AN EVALUATION OF COST METRICS USED WITHIN AERONAUTICAL SYSTEMS CENTER

THESIS

Kevin M. Byrne, Second Lieutenant, USAF

AFIT/GSM/LA/95S-1

DEPARTMENT OF THE AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GSM/LA/95S-1

AN EVALUATION OF COST METRICS USED WITHIN
AERONAUTICAL SYSTEMS CENTER

THESIS

Kevin M. Byrne, Second Lieutenant, USAF

AFIT/GSM/LA/95S-1

The views expressed in this thesis are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

AFIT/GSM/LA/95S-1

AN EVALUATION OF COST METRICS USED WITHIN AERONAUTICAL SYSTEMS CENTER

THESIS

Presented to the Faculty of the School of Logistics and

Acquisition Management

Air Education and Training Command

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Systems Management

Kevin M. Byrne, B.S., M.S.

Second Lieutenant, USAF

September 1995

Approved for public release; distribution unlimited

Preface

With the ongoing emphasis in the Air Force to conform to total quality management concepts, I decided to focus my research on how well we are measuring our progress toward total quality. Specifically, the objective was to determine the effectiveness of Aeronautical Systems Center's (ASC) metrics. The study was limited to cost metrics and they were evaluated using an expert group. The group was asked to evaluate the extent to which metric-driven behaviors contributed to continuous improvement within ASC. Although the evaluation process was difficult, my objective was to provide ASC with information that would improve its progress toward quality through more effective cost metrics.

A great deal of work went into this research. Without the help of generous individuals this thesis would not have been possible. I would like to thank the members of the expert panel for volunteering their time to help with the evaluation. I'd also like to thank Dr. Alan Heminger and Captain Ken Moen for setting up the group support system at Armstrong Labs. In addition, my thesis advisor, Lieutenant Colonel David Murphy, provided the much needed guidance and motivation which was required to complete this project.

2Lt Kevin M. Byrne

Table of Contents

Page

List of Figures

Figure Page

Figure | Page

Figure	Page

List of Tables

Abstract

Aeronautical Systems Center (ASC) system program offices rely on cost metrics to acquire information on program effectiveness and organizational improvement. This study was an attempt to determine the effectiveness of ASC's cost metrics. The effectiveness of the metrics were determined by the desirability of the behaviors which are exhibited when individuals utilize the specific metrics. Using an expert panel to evaluate the metrics, data was collected using the Group Research Laboratory for Logistics located in Armstrong Laboratory at Wright Patterson AFB, OH. Metric diagnostic charts were utilized to arrange the data into a format which would provide effectiveness information.

The results of the study indicated that management should be aware that an integrated approach when utilizing metrics may provide supplementary information about a particular situation. Additionally, management should aim at finding ways to reduce the counter productive behaviors for each metric. Currently, existing metrics are excellent in increasing the likelihood of the behaviors. However, this research shows that a large number of counter productive behaviors are being driven. Consequently, ASC managers must implement policy which will extinguish the undesirable behaviors while keeping the desirable behaviors intact.

Finally, management needs to consider how they are communicating the definitions of the "common" cost metrics. Attention needs to be focused in ensuring that each SPO is using the metrics similarly. After all, common metrics are those that are understood the same way by every individual that uses them.

AN EVALUATION OF COST METRICS USED WITHIN AERONAUTICAL SYSTEMS CENTER

I. Introduction

Background

Quality concepts have altered the operations of many American businesses since the early 1980s. The introduction of quality concepts into business practices was a direct result of consumer dissatisfaction with the quality of many American products. In the past, many firms focused on production with little emphasis given to quality processes. As a result, consumer needs and desires were largely ignored in a firm's quest for profit and market share. As American companies struggled to regain a competitive advantage in world markets, a paradigm shift in basic managerial concepts took place. Barbara Flynn notes this shift toward quality in her article entitled, "Managing for Quality in the U.S. and in Japan":

> In the U.S. during the 1980's, consumers expressed their dissatisfaction with the quality of U.S. goods, particularly in the automobile, consumer electronics, camera, ceramics, textile, and steel industries, where they perceived U.S. quality to be inferior to that of foreign competitors. (Flynn, 1992: 69)

With the turn of the century five years away, the United States has made tremendous strides toward improving products and closing the quality gap that exists between the U.S. and Japan. Civilian and Federal businesses alike are realizing that total

quality involves focusing on all aspects of the process within the organization (Flynn, 1992: 78).

The Federal Government and the Department of Defense (DoD) are currently adopting Total Quality Management (TQM) concepts to "reinvent government." Vice President Al Gore's National Performance Review Initiative is directing that, "federal agencies apply quality management principles to reinvent government" (Federal Quality Institute, 1994: 4). Similarly, the military is being mandated by the DoD comptroller to measure the performance of all government operations (DoD Comptroller, 1992: 1). Judging from these recent policies, TQM concepts are not only pervading the commercial infrastructure of America, but also the government sector as well.

Performance measurement is a fundamental pillar to the success of any organization that utilizes quality concepts. Although many American companies and government agencies have conformed to quality standards, many others experience difficulty in implementing similar standards into their organization. A lack of accurate measurement systems is one reason many good quality programs fail. Research shows that very few American companies measure their performance well. Approximately 29% of American companies evaluate their company's performance less than once a year (Bowles, 1992: 43). Although corporate America is struggling with the concept of measurement in quality programs, the federal government has recognized its importance. The Virginia Productivity Center restates the government's position in their publication Managing Quality and Productivity in Aerospace and Defense

> You cannot manage what you cannot measure. You cannot measure what you cannot operationally define. You cannot operationally define what you do not understand... You will not succeed if you do not manage. (Virginia Productivity Center, 1989: 74)

Aeronautical Systems Center (ASC) system program offices (SPOs) at Wright Patterson AFB rely on measurement of program effectiveness (in terms of cost, schedule, and performance) to evaluate and improve their organizations. They make use of measurement techniques to acquire more information about three specific aspects of their operation: cost, schedule, and performance. The measurement tool program managers rely upon to obtain this information is called a metric. "Metrics, or meaningful measures, are the foundation of an effective measurement system" (Grant and Simpson, 1992: 424). Metrics are useful to program offices because they provide managers with feedback which indicates how well their program is progressing toward established goals (AFSC, 1991).

Problem Statement

Currently, program managers have not determined if their metrics are effective in driving desirable behaviors. In their thesis, An Evaluation of Schedule Metrics Used Within Aeronautical Systems Center, Robert Hayes and Lawrence Miller examined the effectiveness of ASC's schedule metrics. However, the scope of their research did not include cost metrics (Hayes and Miller, 1992). As a result, they recommended further research in the areas of cost and performance metrics. The intent of this thesis is to complement the research performed by Hayes and Miller. Specifically, the objective is to evaluate a set of commonly used cost metrics to determine if they drive behaviors which

contribute to the effectiveness of system program offices in ASC. The results of this thesis will provide upper level managers with an added insight into the effectiveness of their cost metrics. If the results show that the current cost metrics are leading program managers toward behaviors which are detrimental to the overall program, higher level management may want to consider eliminating, or altering, those cost metrics currently in use.

Three research questions will be used to fulfill the objective of this thesis:

1. What are the common cost metrics used within ASC SPOs?
2. What behaviors are driven by the selected cost metrics?
3. Are those behaviors desirable?

Answering these questions will make it possible to accomplish the objective of this research, as well as provide detailed conclusions about the effectiveness (in terms of driving program management behavior) of ASC's commonly used cost metrics.

Scope

This thesis focuses specifically on the cost metrics used by ASC program offices located at Wright Patterson AFB. As a result, the conclusions from this research may only be applicable to those organizations involved with ASC. Although commercial application may be limited, other organizations involved with acquiring weapons systems throughout the military acquisition process may be able to benefit from this thesis.

A final limitation of this thesis relates to the number of metrics used in the research. Metrics are very unique to a program office. As a result, a SPO may have

numerous cost metrics which allows it to evaluate the program's cost status over a specific period of time. Because programs are unique, sometimes the same cost metrics are not found within every program. The metrics evaluated in this thesis are cost metrics that can be found in nearly all of the SPOs in ASC. The purpose of using common cost metrics was to reduce the biased effect that a unique metric (used only in 1 SPO) could have on the applicability of the conclusions and results. Two factors were involved with this decision. First, the expert panel may have difficulty in accurately evaluating a unique metric which they have never utilized. Second, a conclusion may be so specific to the unique metric (and the SPO that utilizes it), that it is not representative of the common cost metrics that are used by all ASC SPOs. As a result, only four cost metrics were evaluated.

Summary

This chapter provided a brief overview of the current emphasis being placed on measurement and metrics as a means to achieving a quality environment. In addition, it identified the lack of research that has been performed in the area of cost metric evaluation in a military environment. This chapter concluded with a discussion of the scope and limitations involved with this research.

Chapter two will review the relevant literature on measurement and metrics, and it will provide a basis for understanding how those concepts engender a quality environment. Chapter three will describe the methodology which was implemented to answer the research questions. Chapter four will present the analysis and results of the research.

Finally, chapter five will conclude the thesis with a discussion of the conclusions and recommendations for further research in the area of metrics.

II. Literature Review

Introduction

Having introduced the concept of measurement in the preceding chapter, this chapter provides a foundation upon which to support the research objective. The literature review covers four general areas. First, metrics are defined, and the literature is used to establish the importance of utilizing measurement systems within a TQM framework. Second, the concept of continuous improvement is discussed. Third, the literature discusses characteristics which produce effective metrics. This section lists those qualities and identifies the attribute which are most important to this research. Finally, the literature review concludes with a brief summary of the prior research that relates to evaluating metrics in a military setting.

Measurement in the TQM Framework

Metrics are tools which give meaning to the inputs of a process so that the status of the system can be understood (AFSC, 1991: 2-1). In addition, metrics provide organizations with a dynamic approach to measurement. Because they can evolve and change in response to an organization's external factors, metrics are an invaluable tool to help managers track changing strategies and goals (Adams, et. al, 1995: 25). Ultimately, an organization's effectiveness can never be reliably measured if it doesn't utilize metrics (Tayntor, 1994: 81).

An effective system of measurement is a prerequisite for every quality program. A 1983 study by David Garvin attempted to show the problems that can result within an organization when a measurement system is not in existence (Flynn, 1992: 75). He performed a study on American companies and found that a large number of firms had no information on defects and failures, and, therefore, had no way to measure organizational performance. As a result, design flaws in products continued to go unnoticed as defective units were sent out to consumers. Because the companies failed to measure, they were not capable of analyzing the status of their processes.

Measurement is a recurring theme within the total quality management literature. Specifically, that theme focuses on the tremendous impact that measurement has on the effectiveness of an organization. In his book, Total Quality Management, author John Oakland states

> Processes operated without measurement are processes about which very little can be known. Conversely, if inputs can be measured and expressed in numbers, then something is known about the process, and control is possible. (Oakland, 1989: 208)

Few have made more contributions to quality than W. Edwards Deming, Joseph Juran, and Philip Crosby. Each has contributed a great deal to the total quality literature, and each has placed a great deal of emphasis on the concept of measurement in their respective theories.

At the conclusion of World War II, Japan's infrastructure was in ruins and its economy was destroyed. W. Edwards Deming is often associated with being one of the designers of Japan's industrial base. His teachings of statistical process control were

aggressively adopted by Japanese companies, which became the basis of their quality philosophy during their growth to become a formidable economic power (Flynn, 1992: 75).

The Deming quality philosophy is based on, "continually monitoring the system process that produces the product or service using statistical process control, and continuously improving that process" (Lyman, 1992: 67). His teachings are organized into a philosophy of management that he calls his 14 points. Deming's fourteen points are listed below in table 2-1:

Table 2-1
Deming's Fourteen Points

1. Create constancy of purpose for the improvement of product and service
2. Adopt the new philosophy
3. Cease Dependence on Mass Inspection
4. End the practice of awarding business on price tag alone
5. Improve constantly and forever the system of production and service
6. Institute training and retraining
7. Institute leadership
8. Drive out fear
9. Break down barriers between staff areas
10. Eliminate slogans, exhortations, and targets for the workforce
11. Eliminate numerical quotas
12. Remove barriers to pride of workmanship
13. Institute a vigorous program of education and retraining
14. Take action to accomplish the transformation (Deming, 1986: 24)

Deming's fifth point is devoted exclusively to measurement, he states, "Improve constantly and forever the system of production and service" (Deming, 1986: 49). By establishing an effective system of measurement, an organization will have the capability to create a baseline, or standard, from which it can track its improvement. In the absence of a standard, organizations cannot improve their production and service capabilities.

Measurement systems are essential to any business, and Deming acknowledges this in his fourteen point management philosophy.

While Deming was in Japan restructuring Japanese management philosophies, Joseph Juran was attempting to initiate the quality movement into the American industrial base. Total Quality Management expert, Joseph Juran is most recognized for his Juran Trilogy. The trilogy consists of quality planning, quality control, and quality improvement. His second concept, quality control, is entirely devoted to measurement. Effective quality control involves three steps. First, an organization is required to evaluate actual performance. Second, the firm must compare actual performance with quality goals. Finally, the differences that exist between its goals and what the organization actually accomplished have to be quantified (Juran 1989: 21).

Juran's philosophy of measurement is focused on the phrase, "say it in numbers". According to Juran, an industrial society increasingly demands higher and higher precision for communicating quality-related information. That higher precision is best attained when we "say it in numbers" (Juran, 1993: 70). To do so requires two concepts, the unit of measure and the sensor. The unit of measure is referred to by Juran as, "a defined amount of some quality feature that permits evaluation of that feature in numbers" (Juran, 1993: 153). For example, acquisition program offices refer to their metrics in terms of cost, schedule, and performance (Dorrell, 1995). If a metric is related to a particular area of cost, that number enables the individual to evaluate the feature by comparing it to past values (or a standard) of the metric. Juran describes the sensor as a detecting device (Juran 1993: 155). Examples of a sensor include technological instruments or reports.

For acquisition offices, the report enables program managers to "detect" changes in cost, schedule, or performance.

Similar to Juran, Philip Crosby also recognizes the importance measurement has in quality operations. Philip Crosby's notion of quality is also based on a system of measurement. He defines quality in terms of conformance or nonconformance to requirements. Crosby states that, "Requirements must be clearly stated so that they cannot be misunderstood. Measurements are then taken continually to determine conformance to those requirements. The nonconformance tested is the absence of quality" (Crosby, 1979: 17). Similar to Deming, Crosby has formulated his own fourteen points which are listed table 2-2.

Table 2-2
Crosby's Fourteen Points

1. Management commitment
2. Quality improvement team
3. Quality measurement
4. Cost of quality evaluation
5. Quality awareness
6. Corrective action
7. Establish an ad hoc committee for the 0 defects program
8. Supervisor training
9. Zero defects day
10. Goal setting
11. Error cause removal
12. Recognition
13. Quality councils
14. Do it over again (Crosby, 132-139)

The third point of his quality improvement program refers to measurement. He states, "It is necessary to determine the status of quality throughout the company. Quality status is recorded to show where improvement is possible, where corrective action is

necessary, and to document actual improvements later on" (Crosby, 1979: 133). Although Crosby admits that a measurement system is only one activity in an interrelated system, he does infer that, without it, quality cannot exist.

W. Edwards Deming, Joseph Juran, and Philip Crosby are known as the pioneers of total quality management. All three have recognized that a lack of a measurement system within an organization will result in its failure to implement quality management principles. The Air Force has embraced the concept of total quality management. As a result, they have established metrics as their system of measurement which will improve processes and increase project performance (AFSC, 1991).

Improper use of metrics, or a poorly designed measurement system, can be detrimental to an organization. Businesses too often use old performance measures to track progress in a changing environment (Adams, 1995: 25). Organizations measure to provide management with insights into: why a system performs the way it does; where improvements can be made; whether the system is in-control or out-of-control; to determine what actions will result in true improvements; and to verify that strategic objectives are being met (AFMC, 1993: 7). When metrics fail to perform these functions, their value in an organizational setting decreases.

Continuous Improvement

Continuous improvement is a basic tenet of TQM (Emmelhainz, 1991: 34). Although the concept of continuous improvement is often associated with modern management methods, its origins can be linked to a 1950s Japanese approach to quality

known as Company Wide Quality Control (Garvin, 1988: 191). This management approach to quality contains four principle elements: the involvement of functions other than manufacturing in quality activities; the participation of employees at all levels; the goal of continuous improvement; and careful attention to the customer's definition of quality (Garvin, 1988: 191). According to David Garvin, the third element refers to continuous improvement.

> Quality programs are expected to aim for perfection, anything less is regarded as an interim goal, to be succeeded by progressively tighter standards. Improvements may be small or may require years of effort, but they will be pursued until defects can no longer be found. (Garvin, 1988: 192)

The ultimate goal and popular catch-phrase is "zero defects." Continuous process improvement is designed to produce zero defects.

The relationship that exists between measurement and continuous process improvement is one of dependence. That is, one cannot occur without the other. Both are vital components to TQM; however, the measurement of inputs, outputs, and processes provide the basis upon which management is able to determine the improvement or deterioration of quality in the system (Oakland, 1989 208).

Air Force literature has defined why it uses metrics to measure performance. According to Air Force Systems Command's Metrics Handbook, "....metrics facilitate quality improvement. Metrics help us understand our processes and their capabilities so we can continually improve them" (AFSC, 1991: 1-3).

Characteristics of Effective Metrics

Air Force Systems Command (now part of Air Force Materiel Command) has defined eight attributes of an effective metric. The attributes are listed below in Table 2-3.

Table 2-3
AFMC Attributes of an Effective Metric

1. It is accepted as meaningful by the customer.
2. It tells how well organizational goals and objectives are being met.
3. It is simple, understandable, logical, and repeatable
4. It shows trend.
5. It is unambiguously defined
6. The data is economical to collect
7. It is timely.
8. It drives the appropriate action or behavior (AFSC, 1991: 2-1).

Because many SPOs create their own metrics to meet changing needs, these attributes provide the evaluation criteria used during the development process. A metric is meaningful if it relates to improvements within the organization. If a metric does not produce information that is useful to those using them, it is not an effective measurement tool. Metrics should also help organizations track internal goals and objectives. Because it is a measurement tool, its purpose is to help the members of an organization visualize how far along they are in accomplishing preestablished goals. Metrics are best implemented if they are logical and understandable to those using them. Without comprehension, they can contribute to unfavorable attitudes for those who are forced to work with them. Metrics should always show a trend, and they should always be unambiguously defined. Trend analysis provides more information because it allows an organization to evaluate performance over a period of time as opposed to a specific moment in time. A metric's end result should mean the same thing for everyone. When

metrics are ambiguously defined they lose some of their power because they are open to many different interpretations. Because information is expensive, the data associated with a specific metric should be economical to obtain. In addition, metrics which provide information after-the-fact are ineffective; they should provide information which relate to future decisions. Finally, a metric must be designed so that it drives an appropriate behavior. The literature suggests that this is the most important attribute of a meaning full metric (Grant and Simpson, 1993: 424).

Metrics Evaluation in Military Acquisition

Although the importance of effective measurement systems is emphasized by Deming, Juran, and Crosby, very little research has been done in evaluating specific metrics in an acquisition setting to see if they possess the characteristics which an effective metric should possess. In 1992, Robert Hayes and Lawrence Miller completed the only evaluation of metrics within ASC which has been performed to date. Their thesis, "An Evaluation of Schedule Metrics Used Within Aeronautical Systems Center" evaluated schedule metrics used within ASC (Hayes and Miller, 1992). Their evaluation was based on whether or not selected schedule metrics drove an appropriate behavior from program managers. Hayes and Miller concluded with eight major findings. Table 2-4 describes those findings.

Table 2-4
Hayes and Miller's Conclusions

1. A single metric may need to be integrated with others to be truly effective.
2. Metrics can lead to sub-optimization in the functional areas within a SPO
3. Behaviors that focus on exploring and improving processes promote continuous improvement. Behaviors that focus on goals, quotas, and the end result usually do not lead to continuous improvement.
4. The field of metrics is a challenging area of study because of the unique features of not-for-profit organizations
5. In order to be fully understood and correctly used, metrics need to be coupled with an objective.
6. If the metric focuses on an activity the SPO has no control over, it shouldn't be used.
7. Too many metrics can be detrimental to the program office.
8. SPOs should consider using Group Support Systems to develop their own internal metrics

Hayes and Miller's work is very valuable in analyzing the effectiveness of schedule metrics. Their research provided the groundwork for research expansion into the areas of cost and performance. This thesis is an extension of the Hayes and Miller thesis, and it focuses on evaluating a set of cost metrics utilized by ASC. At the conclusion of this research, ASC will have had its schedule and cost metrics evaluated in terms of effective behaviors. As a result, performance metrics remain as the last category of metrics to be evaluated within ASC.

Conclusion

The majority of the research within the metrics field relates to how metrics should fit into an organization's strategy and what an effective metric should do. Very little published research exists regarding the evaluation of specific metrics within an

organization. In addition to Hayes and Miller's research, this thesis will expand on the small body of research pertaining to metrics evaluation. Although this research relates specifically to Air Force metrics and ASC, the approach used in this thesis can be duplicated to enhance the body of metric evaluation research.

The objective of this thesis focuses on the behaviors driven by cost metrics. Specifically, the objective is to evaluate a section of commonly used cost metrics to determine if they are driving desired behaviors and contributing to continuous improvement. Air Force Materiel Command states, "For a measure to be meaningful, it must present data that encourages the right action..."(AFMC, 1993: 3). If many of the metrics in use today are leading program managers to act in a way that is detrimental toward organizational effectiveness, the Air Force is not improving its acquisition program offices.

III. Methodology

Introduction

Having previously completed a review of the concepts of measurement and metrics, this chapter addresses how to obtain answers to the research questions posed at the conclusion of chapter one:

1. What are the common cost metrics used within ASC SPOs?
2. What behaviors are driven by the selected cost metrics?
3. Are those behaviors desirable?

A three step methodology was used to provide answers to these research questions. First, using ASC's Executive Information System located in the 88th Communications Group at Wright Patterson Air Force Base, as well as interviews from SPO cost personnel, the common cost metrics used within ASC SPOs were identified. Second, a panel of experts was created to evaluate the metrics. Third, the expert panel participated in a five step evaluation process which made it possible to collect the information needed to create a metric diagnostic chart (MDC). The MDC is a tool which plots the behaviors brainstormed by the expert panel so that the metrics can be analyzed. The chart will be discussed in detail later in the methodology.

Metrics Collection

Before collecting a sample of cost metrics to evaluate, three selection criteria were established. The criterion were used to assure that a meaningful sample of cost metrics

were evaluated in the research. First, the sample had to be representative of the metrics used at most SPOs in ASC. The research will be more applicable to ASC as a whole if common metrics are evaluated. Although evaluating a specific metric which exists in one SPO will be helpful to that organization, it will not provide information into the effectiveness of ASC's cost metrics. As a result, common cost metrics were selected. Second, the candidate cost metrics had to be program level metrics. Metrics exist at all levels of the organization (e.g. functional level, headquarters level, etc.) (Dorrell, 1995). However, the initial decisions that most likely affect the health of a program are made during the early stages of a program at the program level. As a result, program level metrics were selected for this thesis. Third, the selected metrics included a representative sample of the subcategories found within cost. For example, budgeting and cost analysis are the two categories of cost. So, the sample was composed of an equal amount of budgeting metrics and an equal amount of cost analysis metrics. Using these criteria, a representative sample of cost metrics was collected for evaluation.

A sample of cost metrics was collected from ASC. Specifically, the candidate metrics were obtained using the ASC Executive Information System. The system was designed in 1992 so that higher level managers would have access to ASC program office metrics (Dorrell, 1995). The database was initially intended as an on-line system which could collect specific cost, schedule, and performance metrics from all ASC SPOs. However, following 2 years of operation, changes in the requirements of upper level management resulted in the system becoming obsolete (Dorrell, 1995). Although the

system is no longer operational, it provides a standard set of cost metrics which are utilized by all ASC program offices.

The database revealed four cost metrics that are widely used within ASC (Table 3-1). Each metric fit the criteria described above.

Table 3-1
ASC Cost Metrics

Metric 1.	Planned vs. Actual Obligations
Metric 2.	Required vs. Budgeted Funding Levels
Metric 3.	Percent Cost Variance
Metric 4:	Program Cost Disconnect

Instrumentation

Metrics evaluation was performed using a group support system (GSS). A group support system is a tool which can minimize environmental factors in group experiments (Moen, 1995). Group environmental factors include dominance by an individual or group of individuals, free riding by individuals who choose to remain "behind the scenes", and conformance to group pressures. These factors can potentially alter the results of a research effort (Moen 1995). Because an expert panel was used to evaluate the cost metrics, the requirement existed to create an instrument which would accommodate for group dynamics during the evaluation process. As a result, the GSS located in Armstrong Laboratory at Wright Patterson AFB was used for the cost metric evaluation. Robert Hayes and Lawrence Miller discuss the utility of a GSS:

> Group Support Systems are computer-based systems that provide a variety of tools to facilitate the meeting process. In part, these systems are electronic implementations of older methods - e.g. Delphi and Nominal Group Technique - that have been used to improve the quality of meetings over the last 30 years...Over 88% of the users in these studies felt the

system had improved the quality of the decisions reached (Hayes and Miller, 1992: 3-12).

The GSS established a favorable environment from which to create a metric diagnostic chart. The metric diagnostic chart (MDC) is a tool which, "focuses upon the behaviors which will likely result from the measures which are implemented to foster continuous improvement of project performance" (Grant and Simpson, 1992: 424). Specifically, it enabled the metric-driven behaviors to be evaluated in two dimensions -- contribution toward continuous improvement and the likelihood of the behaviors. Appendix A provides a brief description of a MDC. Creating MDCs for this research required a six step process (Grant and Simpson, 1992: 425).

First, an expert panel brainstormed behaviors which were being driven by each of the metrics. During this step, the panel identified as many behaviors as possible. The facilitator encouraged involvement from all group members and very little concern was given to the quality of the behavior presented (Grant and Simpson, 1992: 425). Each group member was seated at their individual computer terminal, and they were required to input a description of the behaviors into their computer The final outcome of this step was an unedited list of unique behaviors provided by all of the group members.

The second step was to have the group refine the list of behaviors by removing ones which were ambiguously defined or redundant. This ensured that the behaviors were appropriately defined, and it allowed synthesis to occur between highly related behaviors (Grant and Simpson, 1992: 425). If any member was unsure of a behavior, the group discussed it and either a better explanation of the behavior was provided or it was deleted

from the list. The importance of this step was to establish a list of behaviors resulting from the metric and not simply outcomes of the metric (Hayes and Miller, 1992: 3-19). The goal of this step was to create a final list of all behaviors which were associated with each metric and to guarantee that all behaviors were understood by the group members.

The third step required the experts to rate the likelihood of the behaviors. This step separated those behaviors which were very likely to occur as a result of using the metric from those which were not very likely to occur. The group was asked the question, "how well does the metric drive the identified behavior"? Each behavior was rated on a 10-point ordinal scale with "10" being an increased likelihood that the behavior would occur and "0" being a decreased likelihood that the behavior would occur as a result of the metric (see Appendix A for more information). At the conclusion of this step, the group had completed one dimension of the MDC.

Step four rated the behaviors according to how well they drove continuous improvement. The group was asked the question, "how well does the metric contribute to continuous improvement"? Again, a 10-point ordinal scale was used with "10" indicating that the behavior was productive in contributing to continuous improvement and "0" indicating that the behavior was counter-productive toward continuous improvement (see Appendix A). With each behavior rated in terms of its likelihood of occurring as a result of the metric and its contribution toward continuous improvement (was it productive or counter-productive toward it), the fifth step created MDCs for each behavior. Using a two dimensional plot, the contribution to continuous improvement was plotted on the x-axis, and the likelihood that the behavior will result from the metric was plotted on the y-

axis (again, refer to Appendix A). From here, all of the behaviors associated with each metric were combined onto one MDC to characterize the effectiveness of each metric.

The final step was to analyze the resulting MDCs. After all relevant behaviors pertaining to each metric were plotted on one MDC, a two part analysis was performed. Part one included a general assessment of the metric and improvement strategies for altering the metric. Part two discussed specific management strategies which could be used to address specific behaviors associated with a metric.

The general assessment within part one of the analysis was used to identify a commonality, or general cluster, of behaviors. This part of the analysis identified trends for each metric. The trends can be determined by the clustering of the behaviors on the MDC. For example, if a metric was counter-productive toward continuous improvement and there was an increased likelihood that the behavior will occur as a result of the metric, the behaviors would cluster in an area that would suggest problems with the metric. Conversely, if a metric was very productive in contributing toward continuous improvement and there is an increased likelihood of the behaviors occurring as a result of the metric, it is likely that the metric is contributing to the overall success of the program. The general assessment was useful for determining if the metrics should be altered (Grant and Simpson, 1992: 428).

Part two of the analysis involved identifying specific management strategies for each behavior. A behavior, or group of behaviors, could be grouped into one of four quadrants: sustain (I), extinguish (II), monitor (III), or cultivate (IV)(refer to figure 3-1). Behaviors which fell into quadrant I were favorable behaviors because they contributed a

high degree toward continuous improvement and they were very likely to occur. Conversely, quadrant II behaviors were undesirable because they were very likely to occur and they were counter-productive in contributing to continuous process improvement. Quadrant III behaviors should be monitored; over time, the behaviors may occur more frequently (i.e. there may be an increased likelihood that the behaviors will occur) and have a counter-productive effect on continuous improvement. Finally, behaviors that were

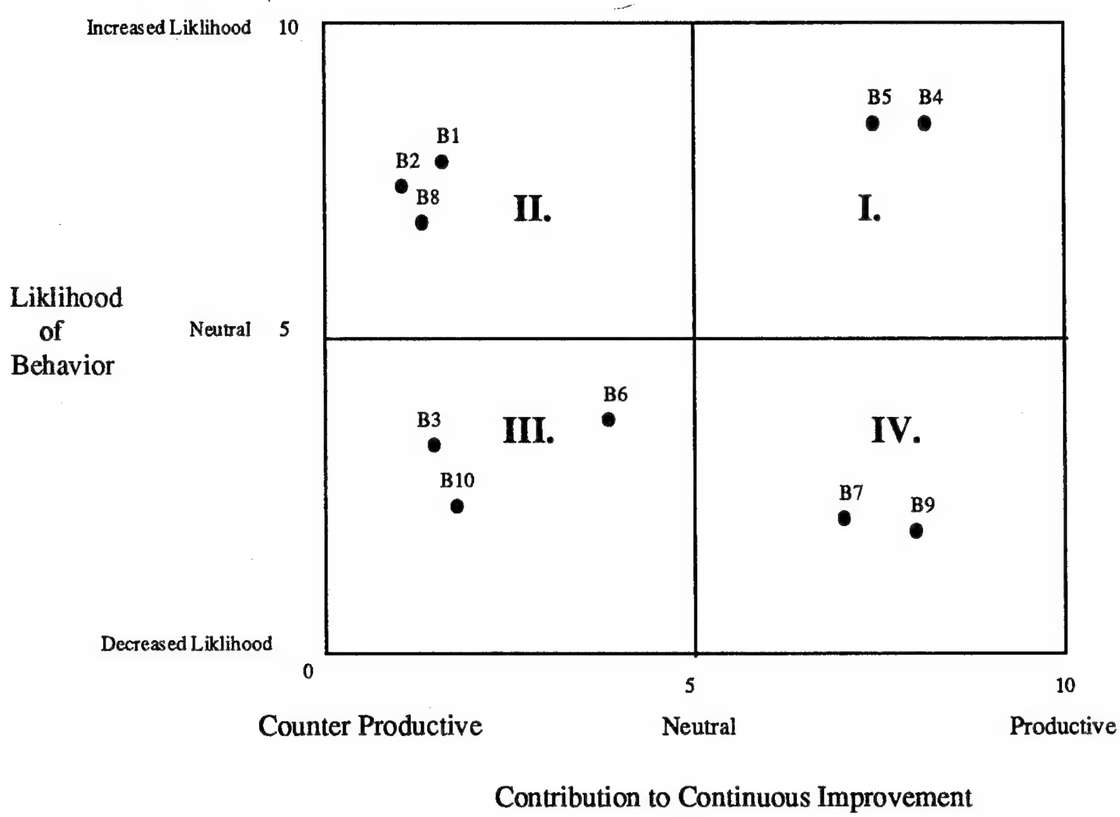

Figure 3-1: Sample MDC

found in quadrant IV were favorable, however, they were not likely to occur. In fact, there is a decreased likelihood that these productive behaviors will occur as a result of the

metric. As a result, a manager would want to cultivate these behaviors so that they were occurring within the organization (Grant and Simpson, 1992: 430). Although these behaviors are productive, the manager's goal is to devise a strategy in which the metric will increase the likelihood of the behaviors.

The outcome of this analysis was to provide the manager with information regarding each specific behavior. With an MDC available, the manager can devise strategies which are capable of attacking a specific behavior or a cluster of behaviors to improve the metric.

Expert Panel

An expert panel was created to perform the first four steps of the MDC creation process. Three criteria were used to ensure that capable, knowledgeable people were chosen to evaluate the selected cost metrics. First, the individuals had to have at least six years acquisition experience at the program level. Six years ensured that the metrics were being evaluated effectively by individuals with "hands-on" experience. Second, members of the panel had to have completed intermediate level program management courses. Intermediate level courses are offered through the acquisition professional development program (APDP). They include Acquisition Planning and Analysis, Intermediate Systems Acquisition Management, and Financial Management in Weapons System Acquisition. Having completed courses of this nature will ensure that the panel members are familiar with the role metrics play in managing an effective program. Although the acquisition professional development program certification (APDP) does provide a measure of how

many formal acquisition training courses have been completed, a certification level was not required of the panel participants. Finally, because this research effort relied heavily on continuous process improvement and measurement concepts, members of the expert panel were required to have had some type of formal TQM training pertaining to metric development or metric implementation.

The size of the group was determined by two factors. The first constraint had to do with the capacity of the group support system (GSS) located in Armstrong Labs at WPAFB. The GSS, has the capability of supporting a group of ten (Armstrong Laboratory, undated). The design of the GSS took into account studies which identified the relationship between group size and group effectiveness. The second factor involved with group size was related to studies on group effectiveness. Paul Hare, author of Creativity in Small Groups, maintains that as group size increases, members become increasingly less satisfied (Hare, 1982: 141). He also states that as group size increases, communication barriers become more prevalent and individuals within the group begin to withdrawal (Hare, 1982: 141). For effective decision making, group sizes of five have been found to be the most effective (Hare, 1982: 142). As a result, five qualified individuals were asked to participate in the metrics evaluation process. Due to scheduling difficulties only four individuals were available at the time of the evaluation.

Conclusion

The results of the metric diagnostic chart analysis provided a basis from which to form conclusions about each metric. The methodology consisted of four phases. Phase

one involved metric identification. Phase two focused on forming an expert evaluation team. Phase three required that the behaviors for each metric were identified. Phase four concluded the methodology by evaluating each metric using a metric diagnostic chart. The resulting MDCs provided a graphic tool from which the effectiveness of each metric could be evaluated. The implementation of this methodology provided answers to the research questions (Table 3-2).

Table 3-2
Thesis Research Questions

1. What are the common cost metrics used within ASC SPOs?
2. What behaviors are driven by the selected cost metrics?
3. Do the behaviors lead to continuous improvement and are they likely to occur?

IV. Metrics Evaluation and Analysis

Four cost metrics were evaluated using the methodology described in the previous chapter. The metrics are listed in table 4-1.

TABLE 4-1
LIST OF METRICS

1. Planned vs. actual obligations
2. Required vs. budgeted
3. Percent cost variance
4. Program cost disconnect

Metric 1. The first metric evaluated was the "planned vs. actual obligations" cost metric. This metric measures actual contract obligations which a program office has allocated versus the planned obligation schedule established at the beginning of the fiscal year. Table 4-2 contains a list of the likely behaviors this metric may drive. The behaviors were collected during the group evaluation process discussed in chapter III.

TABLE 4-2
Likely Behaviors Resulting From Metric 1: Planned Obligations vs. Actual Obligations

1. Causes the SPO to rush into contracts
2. Causes the SPO to plan to obsolete obligation schedule
3. Conduct analysis and /or revision of contract action process
4. Helps coordinate technical definition and contracting efforts
5. Reduces the amount of end of fiscal year contractual actions
6. Forces premature obligation of funds
7. It drives improved planning and forecasting methods
8. Additional work is created to justify needed management actions
9. It causes the SPO to plan for technical definition of contracts
10. It causes the SPO to perform detailed contractor schedule analysis
11. It makes it harder for program managers to handle unforeseen problems
12. May cause obligation of funds based on the schedule of the metric
13. The SPO may offer to return excess funds
14. Estimates of the planned obligations may be high or low based

Rating Summary for Metric 1. Figure 4-1 shows the composite MDC for all of the behaviors relating to Metric 1. Appendix B contains individual graphs for each of the behaviors which were associated with this metric.

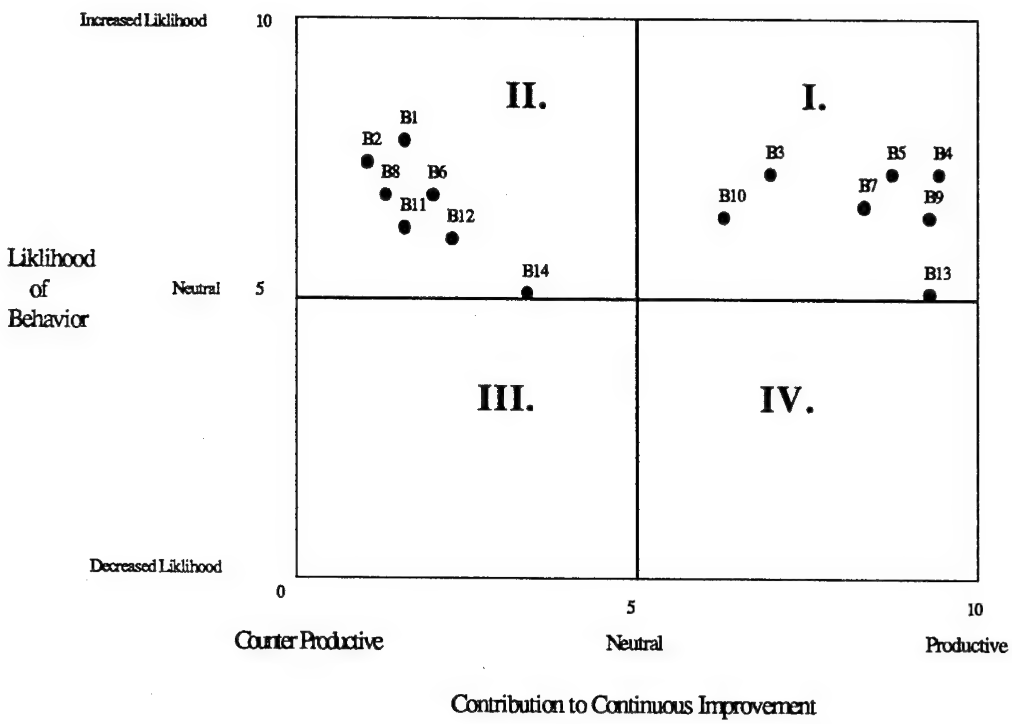

Figure 4-1: Composite MDC for Metric 1

Analysis. The MDC indicates an even split between the behaviors. In terms of contributing toward continuous improvement, half of the behaviors appear to be productive while the other half are counter-productive. Interestingly, very few of the behaviors are found within the "middle ground" of the continuous improvement continuum. This indicates a clear separation in the "good" behaviors and "bad" behaviors

which may be useful in determining the common characteristics of the unfavorable behaviors. If common characteristics can be identified, it could help to focus the modification of the metric to correct for those behaviors.

In terms of likelihood of the metric driving the behavior, the group came to a general consensus. Every behavior was rated in the upper two quadrants, which would indicate an increased likelihood that the behaviors will occur in response to using the metric. Ideally, we would prefer that the counter productive behaviors (B1, B2, B6, B8, B11, B12, and B14) fall below the neutral rating of 5 in terms of likelihood. This would suggest that, although the behaviors are counter-productive, the metric will decrease the likelihood of those behaviors occurring. A metric is more valuable to an organization if it is able to prohibit poor behaviors from occurring.

Seven "unfavorable" behaviors are located in quadrant II of the MDC. Behaviors one, two, six, eight, eleven, twelve, and fourteen are contributing to the negative impact which this metric can have on program offices. Behaviors one, six, and twelve are linked by one common characteristic. These three behaviors seem to be affecting when funds are being obligated. Specifically, they suggest that a project officer will obligate funds for an effort before analyzing the situation in greater detail. For example, if actual obligations are lagging behind planned obligations, a project officer may feel pressured to obligate funds at a greater rate to have a favorable effect on the outcome of the metric. This type of behavior may not be in the best interest of the SPO.

During the evaluation process of this metric, the expert group frequently requested a definition of the metric. Observers of the ongoing evaluation process felt as if each

group member was working with a different definition of the metric. Although each metric chosen for the evaluation was selected for its wide usage within ASC, group members may have been using the metric in a slightly different application. The nature of their jobs dictated how they defined and implemented the metric. As a result, the behavior ratings may have been skewed depending on the perspective each expert adopted in defining the metric.

Metric 2. The second metric evaluated was the, "required vs. budgeted" metric. This metric compares the funds that are required by a program office to those that are actually budgeted for a specific effort. For example, at the beginning of the fiscal year a program will have yearly requirements that it must meet in order to acquire the requested weapon system. However, the SPO may decide not to budget for those requirements until a later date because funding is limited. This metric shows the degree to which funding

TABLE 4-3
Likely Behaviors Resulting From Metric 2: Required vs. Budgeted Funding

1. Brings senior level management attention to programs with big deltas
2. Causes identification of excess funds
3. Causes processing of unfunded requirement request
4. Causes program restructuring
5. Causes request for congressional reprogramming
6. Causes strengthening of budget justification documents
7. Can cause redistribution of funds
8. Cause the SPO to overstate future years requirements
9. Causes the SPO to create contingency list for "fall-out" funds
10. Causes the SPO to lobby harder for budget
11. Causes partial or total contract termination
12. Causes managers to violate appropriation law
13. Causes procurement of unauthorized goods

levels lag behind requirements or visa versa. This information is useful for planning future funding requests. Table 4-3 contains a list of the likely behaviors this metric may drive.

Rating Summary for Metric 2. Figure 4-2 shows the composite MDC for all of the behaviors relating to Metric 2. Appendix B contains individual graphs for each of the behaviors examined during the evaluation

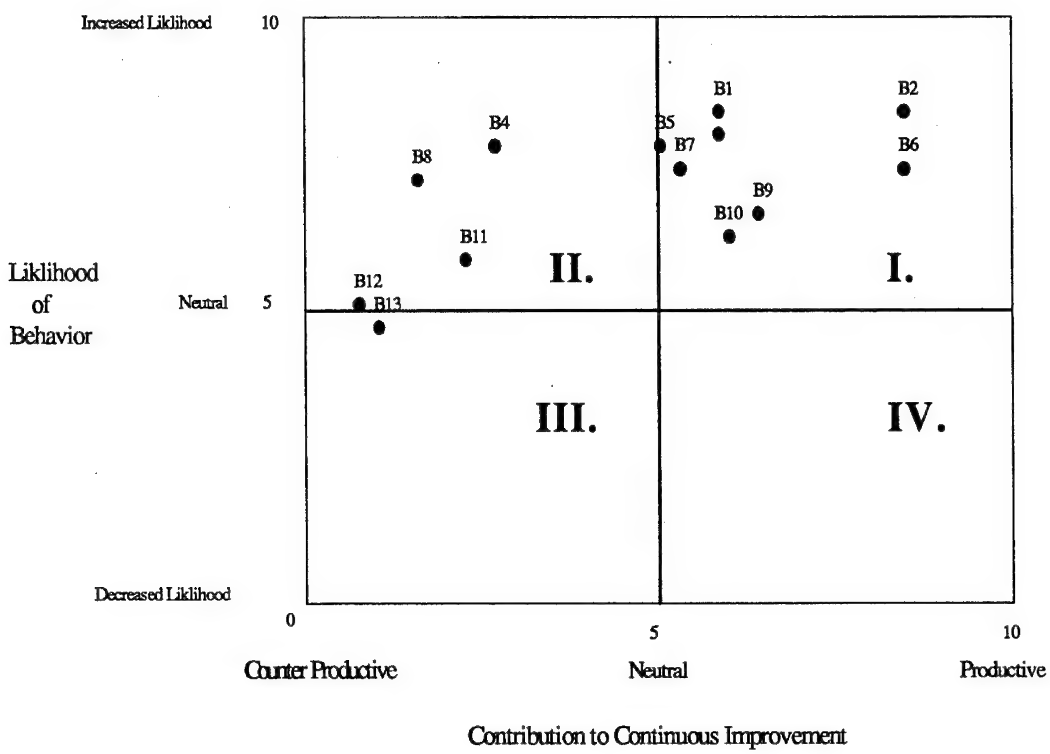

Figure 4-2: Composite MDC for Metric 2

Analysis The MDC for metric 2 indicates a slight uneven split between the behaviors. It appears that three general clusters of behaviors have formed. The first group is clustered within quadrant II. The second group is tightly bunched to the left of

quadrant I, and the third group (consisting of B2 and B6) lies to the right of quadrant I. In terms of contributing toward continuous improvement, more than half of the behaviors appear to be productive while the remaining are counter-productive. In terms of likelihood, the metric is increasing the likelihood that the behaviors will occur. With the exception of B12 and B13, the remaining behaviors lay well above a neutral rating of 5.

Behaviors four, eight, eleven, twelve, and thirteen were located within the first cluster of behaviors, however, only three of these (B4, B8, And B11) were likely to occur. Because the likelihood of behaviors twelve and thirteen were rated neural or below, these were not as significant as the remaining three behaviors. In fact, behavior thirteen may be beneficial toward a program office because it is decreasing the likelihood of a counter-productive behavior occurring. The characteristics associated with the unfavorable behaviors deal primarily with requirements "fudging". The project officer may be tempted to overstate program requirements if it results in greater funding. This results in creating a distorted view of the program. The metric can be improved if it can be altered to limit its "fudging" effect on project officers.

Overall, the metric is effective in driving favorable behaviors. Although five behaviors were located within quadrant II, two of the behaviors (B12 and B13) have minimal influence on the effectiveness of the metric due to their neutral likelihood rating. Management should consider focusing its attention on the "fudging" behaviors which are being driven by the metric. The metric will be slightly more effective if management is able to shift the likelihood of these behaviors down in the MDC. Although the current location of the unfavorable behaviors is not alarming, management may want to consider

focusing its attention on minimizing the likelihood of these behaviors should they get out of control.

Metric 3. The third metric evaluated was the, "percent cost variance" metric. This metric looks at the cost status of the program by taking the actual work that has been performed and comparing it to the budgeted amount of work that was planned to have been completed at that point in time. The metric is given as a percentage. A positive percentage suggests that the contractor is under budget, and a negative percentage indicates that the contractor is over budget. Table 4-4 contains a list of the likely behaviors this metric may drive.

TABLE 4-4

Likely Behaviors Resulting From Metric 3: Percent Cost Variance

1. Causes the contractor to explain the variance
2. Creates additional work to justify deviations
3. Causes the contractor to apply management reserve to overruns
4. compare contract overrun with budget status
5. Causes increased oversight of contractors
6. Causes the SPO to look into variances
7. Promotes cost cutting measures
8. Budget for anticipated overruns
9. Helps to control contractor spending
10. Causes the SPO to rebaseline
11. Helps to keep programs under control
12. Causes examination of schedule performance
13. Causes the SPO to watch overhead rates
14. Causes the SPO to mandate cost cutting
15. Revise billing, liquidation, and profit rates
16. Overstating the budgeted cost
17. Causes cancellation of programs
18. Helps generate new cost estimating techniques
19. Causes the contractor to cut corners
20. Report contractor performance in CPAS
21. Indicates the need for better cost estimating techniques
22. Biased methods for contractor taking credit for work performed

Rating Summary for Metric 3. Figure 4-3 shows the composite MDC for all of the behaviors relating to Metric 3. Appendix B contains individual graphs for each of the behaviors examined during the evaluation

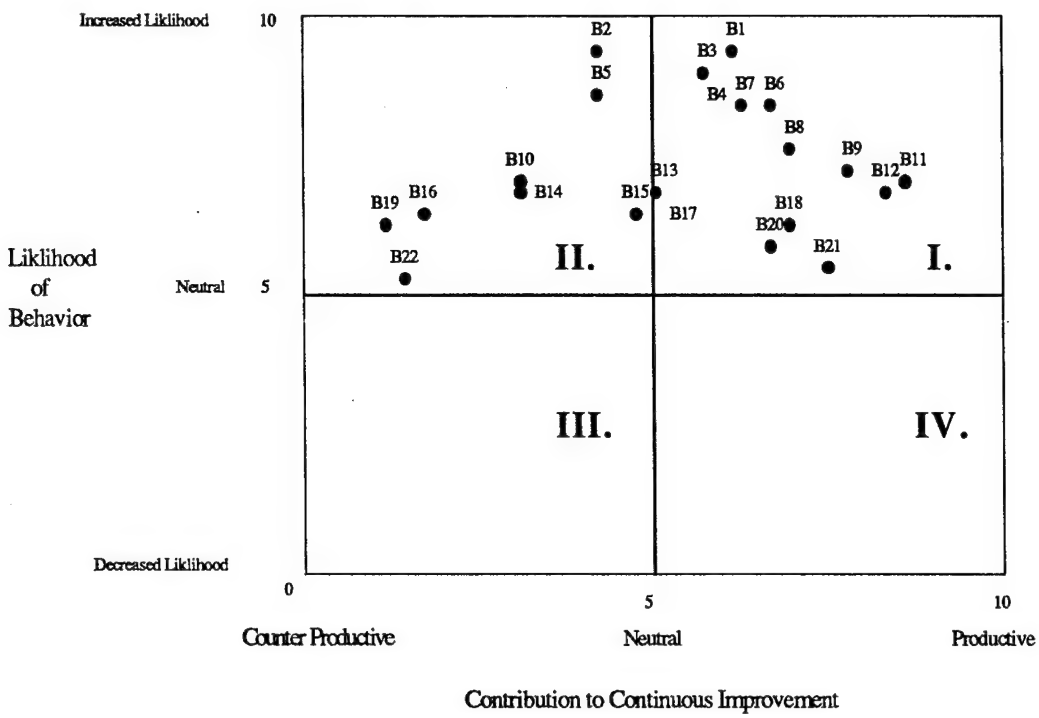

Figure 4-3: Composite MDC for Metric 3

Analysis. Similar to metric 2, the majority of the behaviors were found in quadrant I of the MDC which indicates that the metric is very effective in driving desirable behaviors. Although eight behaviors were located in quadrant II, two behaviors stood out. Behaviors two and five were rated as very likely to occur, however, they weren't rated as extreme counter productive behaviors. If these behaviors shift toward the left over time, they could become very problematic for a program office. The behaviors were

related to increasing the oversight of the contractors and increasing the work load of the project officers to justify cost variances. The expert panel suggested that cost variances sometimes can't be explained. However, when the metric shows a negative variance, the project officer feels compelled to find a reason even if one doesn't exist. As a result, the project officer increases his/her oversight on the contractor in hopes of explaining the variance.

A second area of concern lies in the cluster of behaviors which includes B16, B19, and B22. Currently, they are positioned very close to the neutral rating in terms of likelihood. However, they are very counter productive behaviors, and if they shift upward, they could potentially decrease the effectiveness of a program office.

Overall, a slight majority of the behaviors being driven by the metric are favorable, which indicates that the metric is an effective metric. However, it is not a perfect metric. Ideally, we would like the behaviors in quadrant II to either shift down into quadrant III or move to the right to enter quadrant I. Currently, the metric doesn't drive any behaviors which require immediate attention. However, similar to each metric in this research, a future evaluation is recommended to analyze the movement of the unfavorable behaviors.

Metric 4. The fourth metric evaluated was the, "program cost disconnect" metric. This metric measures the program's budgeted multiyear funding for development, production , and operations and maintenance versus the program's identified funding requirements. The benchmark is a zero cost disconnect in each of these types of funding. This metric is similar to metric 2 in that it compares required versus budgeted funding,

however, it gives the project officer a different perspective. Metric 2 is concerned with only one aspect of the project (e.g. development, production, or operations and maintenance individually). However, metric 4 shows the status of the operations in terms of how development, production, and operations and maintenance compare with each other. Table 4-5 contains a list of the likely behaviors this metric may drive.

TABLE 4-5

Likely Behaviors Resulting From Metric 4: Program Cost Disconnect

1. Add or delete content, speed or slow schedule, as required
2. restructure program plan to match budget
3. Causes the SPO to carefully plan development., production, and O&M funding
4. Causes speculation on political environment--wasted effort
5. Causes the SPO to request more funds
6. Provides justification for increased funding
7. Causes the SPO to request reprogramming of money
8. Cause revisions to requirements process
9. Causes reduction in number of programs (fund fewer better)
10. request direction for use of excess funds

Rating Summary for Metric 4. Figure 4-4 shows the composite MDC for all of the behaviors relating to Metric 4. Appendix B contains individual graphs for each of the behaviors examined during the evaluation

Analysis. Very few unfavorable behaviors originated from this metric. Four behaviors were identified in quadrant II, however, three of the four are located very close to a neutral rating in terms of their contribution toward continuous improvement. This suggests that they really don't detract from the effectiveness of the organization. They are

counter productive behaviors, but not enough to have a considerable effect on the program office.

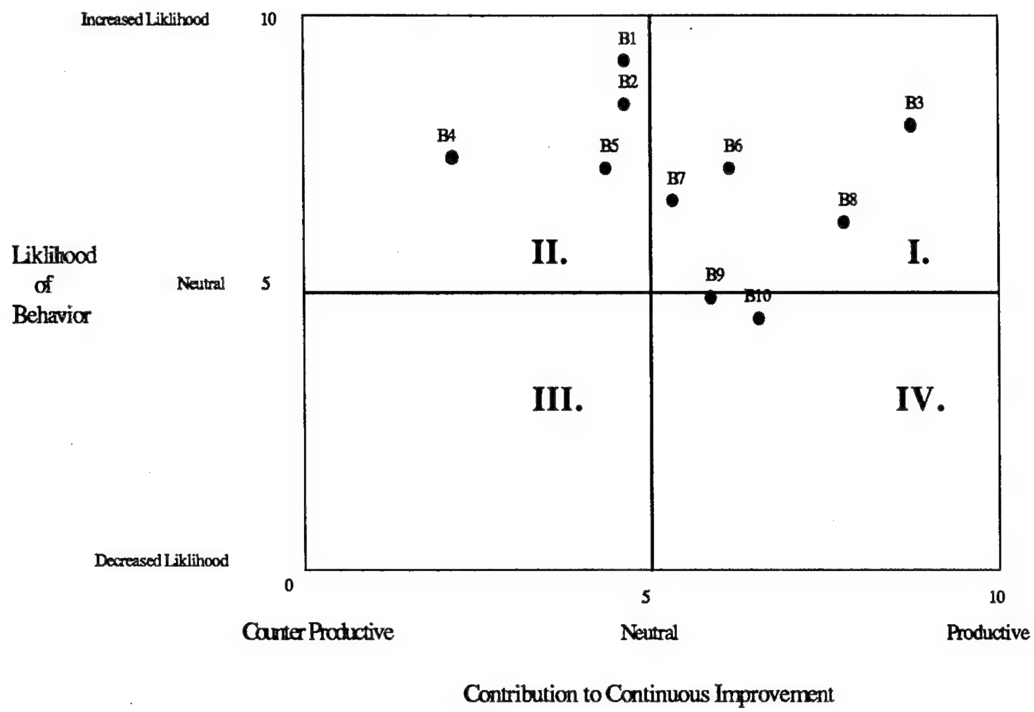

Figure 4-4: Composite MDC for Metric 4

Behavior four is the only one which could be termed a "problem behavior". It relates to the highly political environment associated with the acquisition business, and it suggests that program managers may become overly concerned about its effects on the program as a consequence of the metric. This is a concern, however, modifying a metric to limit this single behavior may be very difficult - if not impossible. After all, a political environment pervades the military acquisition business, and project officers are expected to encourage political support for their program so that funding levels remain consistent

each year. Although too much emphasis in the political arena is harmful, it may be difficult to adjust this metric to account for one behavior.

Unlike the behaviors driven by the other metrics, this was the first metric which drove behaviors in quadrant four. Behaviors within this area are characterized as productive, however, the metric is decreasing the likelihood that these productive behaviors will occur. As a result, this metric is slightly suppressing favorable behaviors which could lead to continuous improvement within the program office. Ordinarily, this would decrease the effectiveness of the metric, however, the position of the behaviors very close to the neutral rating ensures that they will not detract from its ability to improve the SPO.

In addition to the unfavorable behaviors, the group found it very difficult to separate this metric from metric two. Although the metrics provide unique information, the panel argued that the metrics could not be evaluated separately. Because they both focus on required and budgeted funding levels, many of the behaviors which the metrics drive are similar. Consequently, the group struggled to brainstorm a representative list of behaviors. The group suggested that the two metrics provide more information when they are integrated as opposed to when they are utilized individually.

Overall, metric four is effective. Although unfavorable behaviors were found within the MDC, the metric is very weak in driving those behaviors. As a result, the metric should remain unchanged.

Conclusions.

This analysis provided the foundation which leads to four major conclusions of this thesis. The conclusions attempt to synthesize the information found within this chapter in order to guide management's action to improve ASC's cost metrics. The conclusions and recommendations relating to specific behaviors within this chapter are meant to act as a guide for managers who utilize these cost metrics. The flexible nature of metrics allow managers to utilize originality and creativeness to reduce the likelihood of counter productive behaviors.

V. Conclusions and Recommendations

Conclusions

Four major conclusions resulted from this thesis effort. They are listed below in Table 5-1.

Table 5-1
Conclusions

1. ASC's cost metrics are effective in driving behaviors.
2. Integrated metrics provide more information than a single metric on its own.
3. ASC's common cost metrics are understood differently throughout ASC. The same metric can have many different meanings.
4. An equal amount of counter productive behaviors resulted from every productive behavior.

ASC's cost metrics are driving a response from project managers. Referring to the composite MDCs in chapter IV, one is able to identify a trend in where the behaviors were located. Nearly, all of the behaviors were located in quadrants I and II. Again, quadrant I behaviors are favorable because they contribute a high degree toward continuous improvement. Conversely, quadrant II behaviors are undesirable because they are counter productive behaviors which are very likely to occur. ASC's cost metrics are effective in driving behaviors. Although not all of the behaviors are favorable, we see that the metrics are eliciting a response as a result of the metric. It is important to note that these cost metrics are stimulating project managers toward behaviors. Suppose that the panel had rated the likelihood of the behaviors so that they clustered around a neutral rating of 5 on the MDCs. This would suggest that the metrics were completely ineffective in stimulating an action or behavior, and Air Force Materiel Command has stated that its metrics will

drive an appropriate action or behavior. This wouldn't have been the case had the metrics been clustered around a neutral likelihood rating. Conversely, if the behaviors had clustered around a decreased likelihood, this would indicate that the metrics are reducing the likelihood of the behaviors occurring. This would be favorable if the behaviors were counter productive and unfavorable if they were productive. As it turned out, very few of the behaviors were located within this decreased likelihood range. Ultimately, we would like the behaviors found within quadrant II to shift down so that there would be a decreased likelihood of the counter productive behaviors within the SPO. Although not all of the behaviors relating to these metrics are favorable, the metrics are generating a response from program mangers.

In some cases, integrated metrics may provide more information than a single metric. This conclusion supports a similar finding by Hayes and Miller regarding schedule metrics (Hayes and Miller, 1992: 5-1). Of the four cost metrics evaluated in this thesis, two directly complement each other. The required vs. budgeted and the program cost disconnect metric provide more information when integrated than they do when they are observed individually.

Both metrics deal with required and budgeted amounts, however, the program cost disconnect metric is more concerned with how development, production, and operations and maintenance funding levels compare with each other. Unlike the broad, program-wide perspective of the required vs. budgeted metric, the program cost disconnect metric attempts to give the user a more defined picture of, "where the money is going". Conversely, the required vs. budgeted cost metric combines the development, production,

and operations and maintenance funding into one category. Although they can be utilized separately, an integrated approach to using the metrics will provide the user with a more detailed status of the program.

A single metric rarely stands on its own. A method of cross checking metrics is useful to gain an accurate picture of an organization's progress toward its goals. For example, if actual obligations are lagging behind planned obligations, the required versus budgeted cost metric could explain why. A SPO may not be budgeting enough of its resources to meet the planned obligations schedule. By integrating and cross-checking metrics an organization is able to explain unfavorable situations and, more importantly, identify which areas are causing the problems.

Metrics are very powerful measurement tools because they can be created and implemented in many diverse environments. However, specific metrics are designed with a common use in mind. The four cost metrics analyzed in this research were developed by ASC because they provided information which could be compared across SPOs. Organizations were expected to utilize the metrics in the same way, using the same definitions. During the evaluation process of this thesis, the researcher encountered difficulty in ensuring that everyone had the same working definition of the metrics. Often, panel members were unsure if their working definition of the metric was congruent with the definition of what other members were using. At the beginning of the evaluation process each metric was defined in order to ensure that they would be rated based on one shared meaning. Because the metrics were common cost metrics used in every SPO within ASC, the research group didn't anticipate a large disagreement in the definitions

between panel members. However, during the evaluation, expert panel members went into extensive discussion regarding the definitions of the metrics. The experts expressed their concern that the definitions which were provided were not the same definitions that they were using "in the field". The high variance in behavior ratings is evidence that the group may have rated the metrics based on their prior meanings.

The fact that an expert panel had to come to a consensus regarding "common" cost metrics was interesting. It is an indication that ASC's "common" cost metrics are not providing consistent information which can be used as a comparison across SPOs. If these metrics truly are basic cost metrics which every SPO should be using, everyone should be utilizing them the same way. Observations during the metric evaluation process failed to support that statement. The high variability in the rating could a result of a lack of understanding between group members. Further discussion and understanding of the metrics by the group members may have reduced this variance.

A final conclusion, which is also supported by the work performed by Hayes and Miller, relates to a neutralizing effect found within the MDCs. Referring back to the composite MDCs found in chapter IV, it becomes apparent that an equal number of productive behaviors are neutralizing the counter productive behaviors. This neutralizing effect can also be seen in the Hayes and Miller thesis regarding schedule metrics. Their composite MDCs showed that nearly all of their behaviors were found in quadrants I and II indicating that ASC's schedule metrics are effective in driving behaviors, however, they are split between driving productive and counter productive behaviors. This split is not an exact one-for-one split. A trend does exist which slightly favors the productive behaviors.

This trend cannot be supported by Hayes and Miller's research, which suggests that ASC's cost metrics are driving more productive behaviors than its schedule metrics.

Overall, ASC's cost metrics are effective. Management should be aware that an integrated approach when utilizing metric two and metric four may provide supplementary information about a particular situation. Additionally, management should aim at finding ways to reduce the counter productive behaviors for each metric. Currently, the metrics are excellent in increasing the likelihood of the behaviors. However, this research shows that a large number of counter productive behaviors are being driven. Although the counter productive behaviors for each metric are not alarming, they may reach that stage in the future. A solution to decreasing these behaviors may be as simple as recognizing that they exist and communicating their existence among the individuals that use the metrics. Finally, management needs to consider how they are communicating the definitions of the "common" cost metrics. Attention needs to be focused in ensuring that each SPO is using the metrics similarly. After all, common metrics are those that are understood the same way by every individual that uses them.

Recommendations For Further Research

Chapter two of this thesis described the importance of having effective measurement systems in place to continually improve an organization. Metrics are a tool which can provide an organization with a capable measurement system. Certainly, they have provided the foundation in the Air Force's drive to implement quality concepts. Four recommendations for further research are outlined in this section which will improve the

Air Force's ability to use and improve their current use of metrics. First, the final category of metrics should be evaluated within ASC. Second, metrics in other commands should be evaluated. Third, ASC metrics should be tested to determine if they meet the remaining AFMC characteristics of effective metrics. Finally, a two part research project regarding the Group Support System and the metric diagnostic chart would provide Air Force organizations the ability to evaluate their own metrics.

Three types of metrics exist within the Air Force metrics structure. They are cost, schedule, and performance metrics. In their thesis, An Evaluation of Cost Metrics Used Within Aeronautical Systems Center, Hayes and Miller studied the effectiveness of schedule metrics. Using a similar methodology, this thesis focused on the effectiveness of cost metrics used within Aeronautical Systems Center. Performance metrics remain as the final category of metrics to evaluate, and this methodology could provide the basis for a thesis effort which could evaluate a common set of performance metrics used within ASC. Once the three types of metrics have been evaluated, an additional study could be performed to integrate the results and establish conclusions regarding the overall effectiveness of ASC's metrics.

An additional recommendation relating to metrics evaluation would be to perform research regarding the effectiveness of the metrics in other commands. For example, it may be interesting to determine the effectiveness of Air Force's operational metrics. Do sortie rate metrics lead to faulty repair work? Are downtime metrics leading crew chiefs to bypass approved supply channels? An evaluation of these types of metrics found with Air Force operational units would be useful in determining what types of behaviors these

metrics are driving. More importantly, it would be useful in determining if those behaviors are contributing to organizational improvement. Although Air Force Systems Acquisition is an important aspect of the Air Force's overall capability, one could argue that operational metrics are more important. If operational metrics are driving undesirable behaviors it could be harmful to our warfighting capability as well as for those operating and maintaining the various weapons systems.

Chapter II of this thesis discussed the attributes of an effective metric. Although this thesis focused on evaluating the metrics in terms of behavior being driven, seven other AFMC attributes remain to be tested. ASC's cost, schedule, and performance metrics could be evaluated according to the seven other AFMC criteria that are listed in chapter II. The results of this type of a thesis effort would provide a more in-depth analysis regarding the relationships of ASC's metrics. For example, maybe cost metrics are better at showing trend than schedule metrics. This type of research would provide answers to those types of questions.

Finally, the metric diagnostic chart and the group support system were essential tools in collecting and analyzing the data involved with this research. The group support system made it possible to collect data in an organized manner, and the MDC provided the capability to order that data in a meaningful way. A way in which the effectiveness of the metrics could be determined. The metric diagnostic chart (MDC) is a relatively new tool. Consequently, very little is known about its ability to be adapted to many different situations. For example, although this thesis was very similar to the Hayes and Miller thesis regarding schedule metrics, the scales on the MDCs here were altered slightly. On

both scales, they used a continuum ranging from very poorly (0) to very well (10). The scales here range from productive to counter productive in terms of contribution to continuous improvement and from an increased likelihood to a decreased likelihood in terms of the behavior's likelihood. Although both scales provided for a complete evaluation of the metrics, each provided a different type of information. So, the MDC is a very flexible tool. As a result, it can be tailored to meet the needs of all types of organizations interested in evaluating their metrics. A research effort which described the capabilities and limitations of these metric evaluation tools, as well as others, would be useful to other Air Force organizations.

Summary

The development and utilization of metrics within the Air Force is helping organizations improve its managerial processes. The conclusions from this research will enhance ASC's ability to effectively measure costs within government acquisition offices, and it will improve the Air Force's ability to create and implement more effective metrics which drive favorable behaviors. Reducing counter productive behaviors, integrating metrics, and providing a collective understanding of common metrics will ensure that this occurs.

Appendix A. Metric Diagnostic Chart Description

The metric diagnostic chart (MDC) is a tool that, "focuses upon the behaviors which will likely result from the measures which are implemented to foster continuous improvement of project performance" (Grant and Simpson, 1992: 424). Figure A-1 shows a sample MDC. The dots in the chart (labeled B1-B10) represent individual behaviors which were brainstormed during the expert panel evaluation.

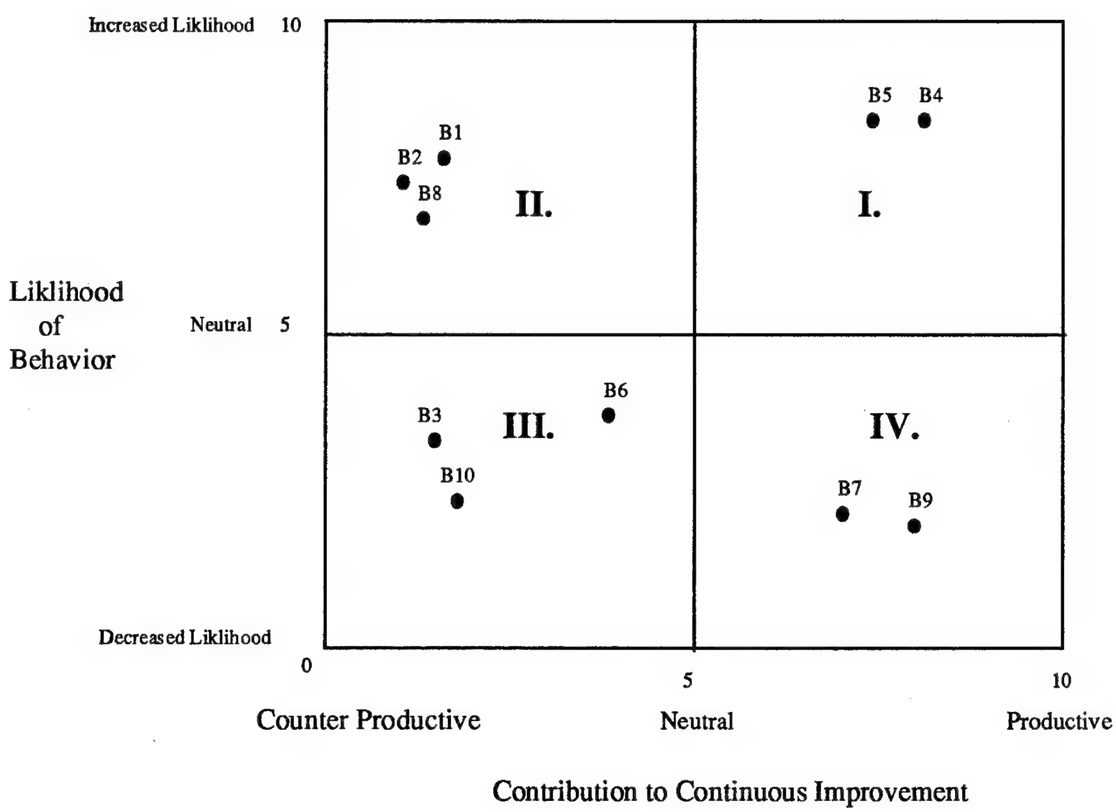

Figure A-1: Sample MDC

The chart contains two scales. The x-axis scale rates the behaviors in terms of their effect on continuous improvement to the organization. In other words, do the behaviors lead an organization to continually improve their processes? The scale ranges

from 0-10. A rating of 0 would indicate that the behavior is counter-productive to an organizations progress toward continuous improvement. A 5 would indicate that the behavior is neutral, and a rating of 10 would show that the behavior is productive in driving an organization to continually improve.

The y-axis rates the behavior is terms of how likely it is to occur. For example, suppose the expert panel brainstormed the following behavior: program managers embezzle money. If there is an increased likelihood of this behavior occurring as a result of the metric, the behavior would receive a 10 by the expert evaluators. However, if the metric decreases the likelihood of the behavior occurring, the behavior would receive a 0. Again, a rating of 5 indicates that using the metric neither increases nor decreases the likelihood of the behavior from occurring. Both scales are continuous scales. As a result, ratings can range anywhere from 0-10.

Appendix B. Individual Behavior Ratings

Metric 1: Planned vs. Actual Obligations

Metric 1 Behavior 1. Causes the SPO to rush into contracts.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	1	2	1	-	8.00
How well does the behavior contribute to CI?	1	1	2	-	-	-	-	-	-	-	2.25

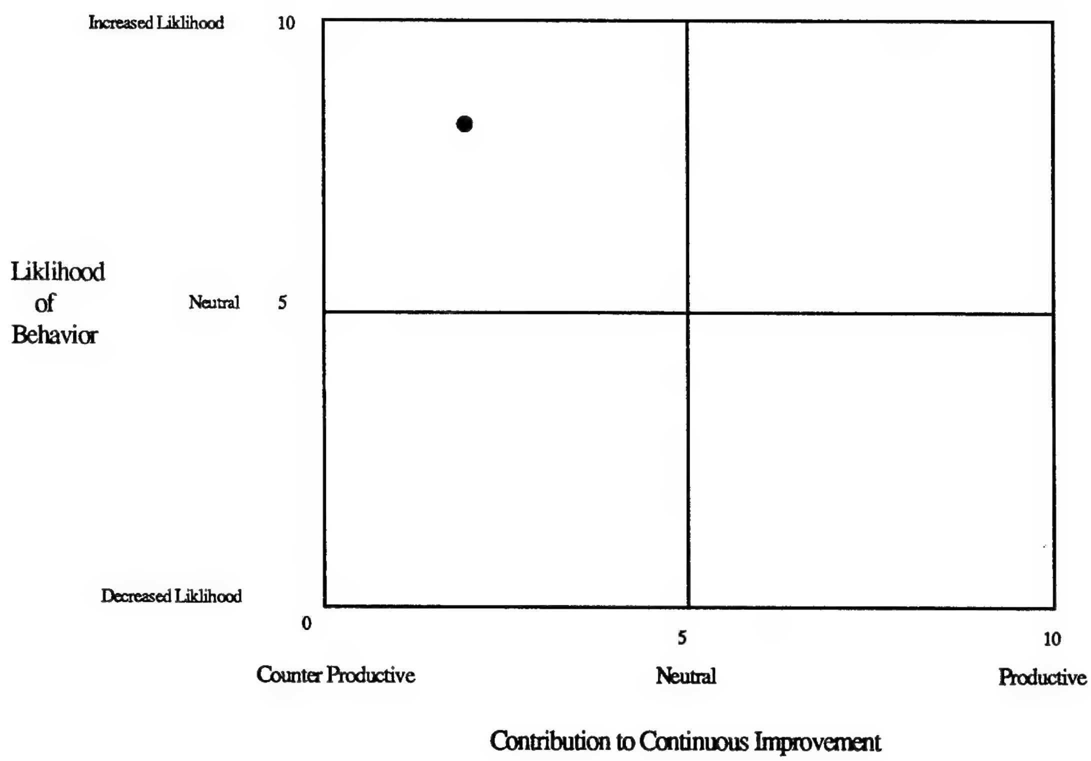

Figure A-1. Metric One, Behavior One

Behavior 2. Causes the SPO to plan to obsolete obligation schedule.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	-	-	-	1	1	7.75
How well does the behavior contribute to CI?	2	1	1	-	-	-	-	-	-	-	1.75

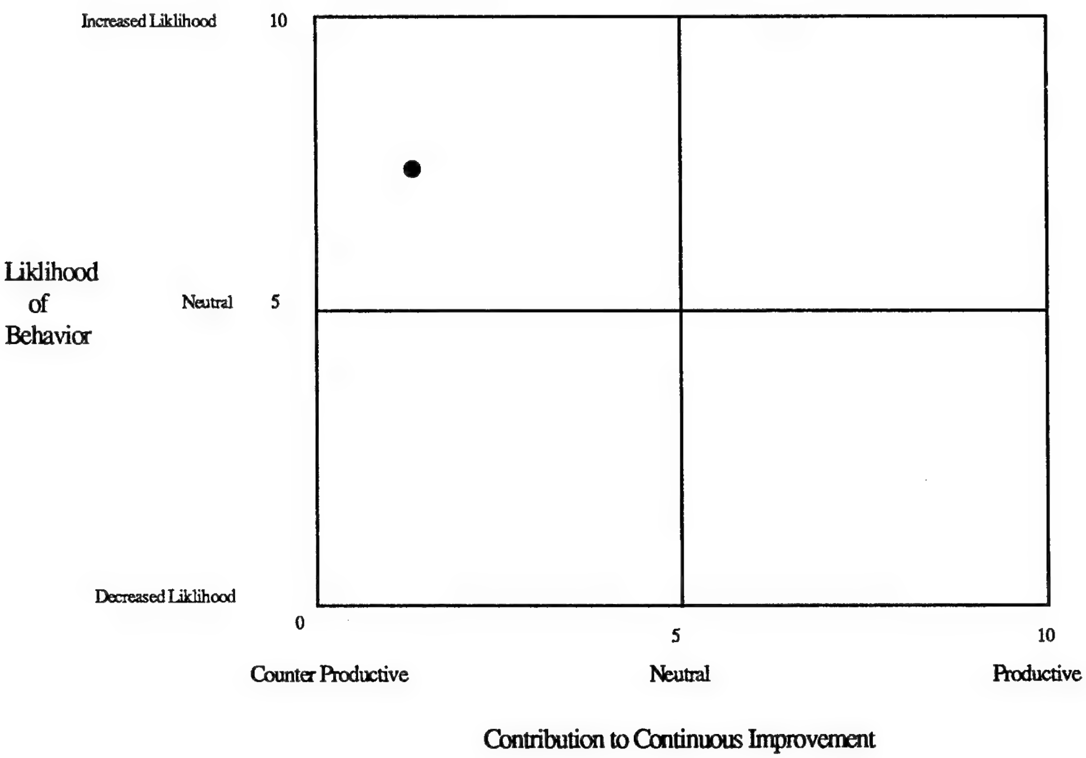

Figure A-2. Metric One, Behavior Two

Behavior 3. Conduct analysis and/or revision of contract action process.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	2	2	-	-	7.50
How well does the behavior contribute to CI?	-	-	-	-	-	-	2	1	1	-	7.75

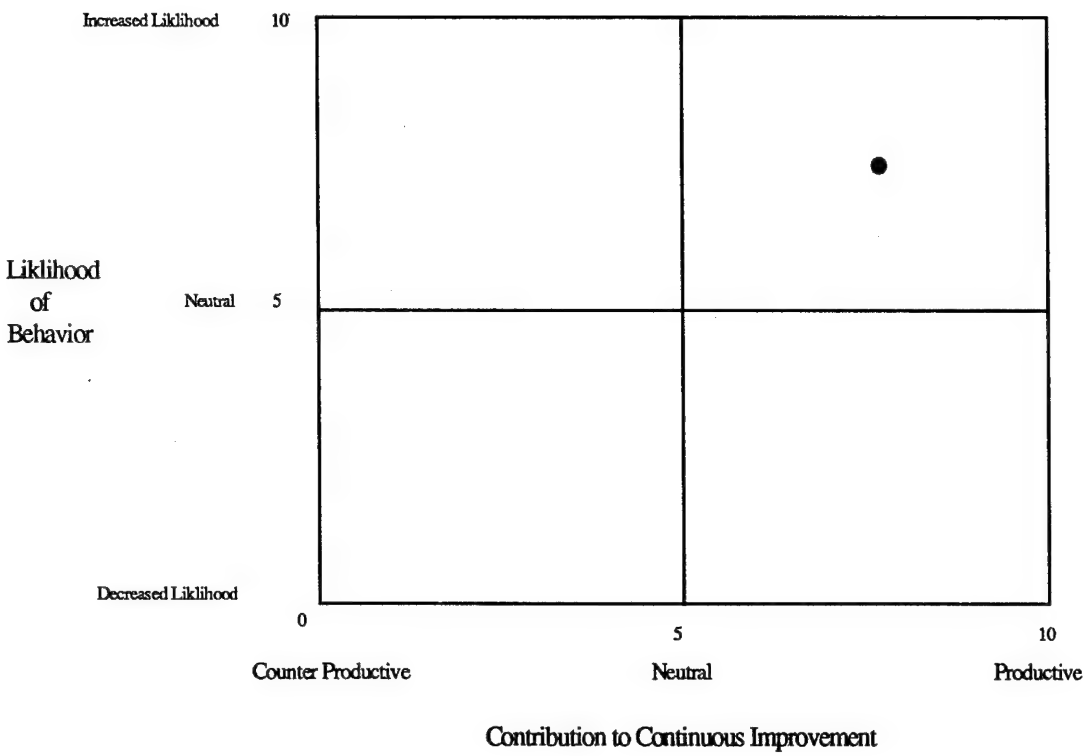

Figure A-3. Metric One, Behavior Three

Behavior 4. Helps coordinate technical definition and contracting efforts.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	1	1	1	1	-	7.50
How well does the behavior contribute to CI?	-	-	-	-	-	-	-	1	1	2	9.25

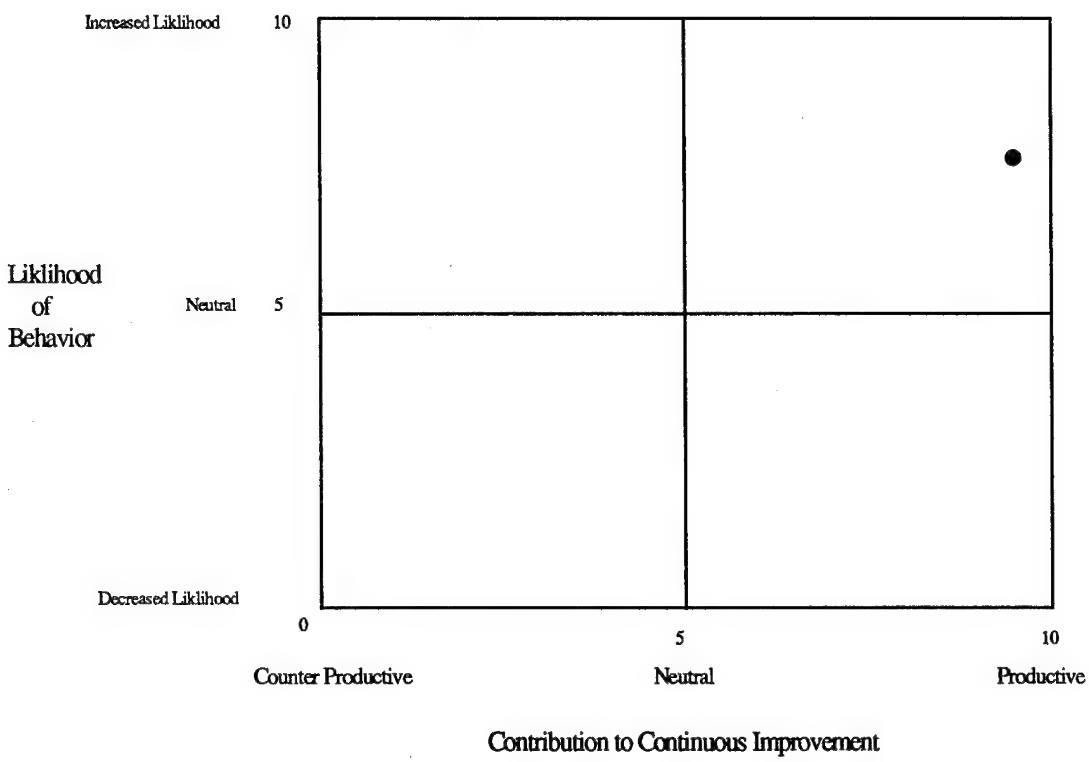

Figure A-4. Metric One, Behavior Four

Behavior 5. Reduces amount of end of fiscal year contractual actions.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	-	1	-	2	-	7.50
How well does the behavior contribute to CI?	-	-	-	-	-	1	-	-	3	-	8.25

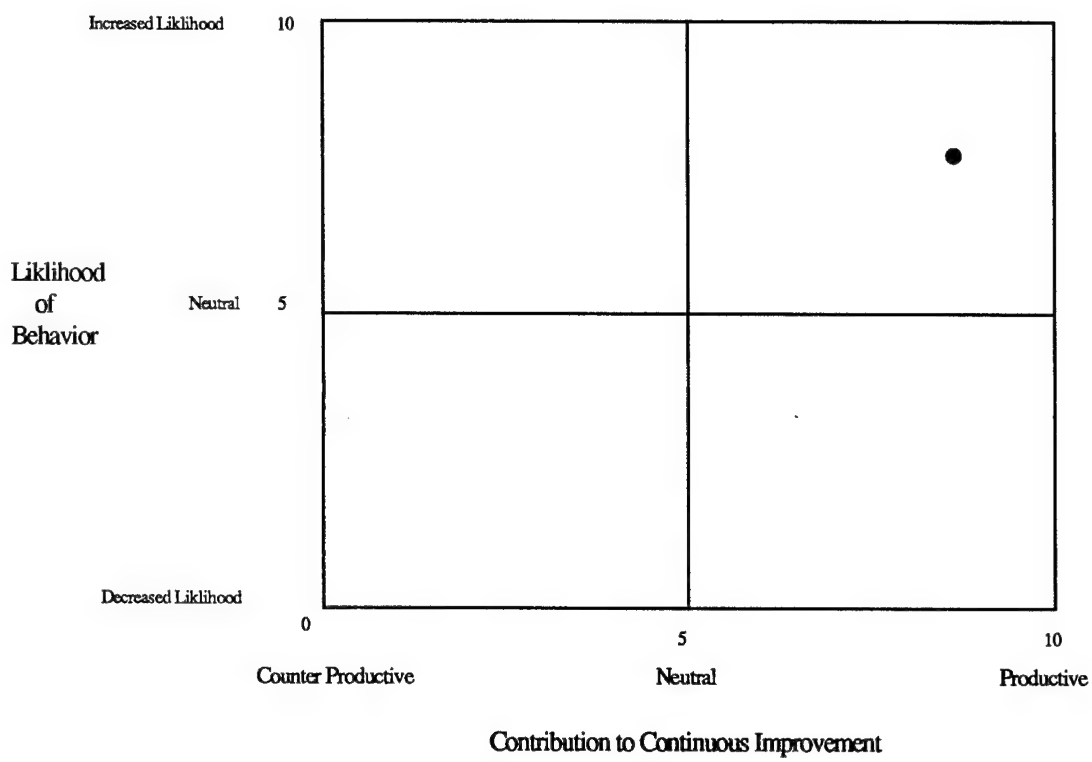

Figure A-5. Metric One, Behavior Five

Behavior 6. Forces premature obligation of funds.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	1	-	-	-	2	1	-	7.25
How well does the behavior contribute to CI?	1	-	2	1	-	-	-	-	-	-	2.75

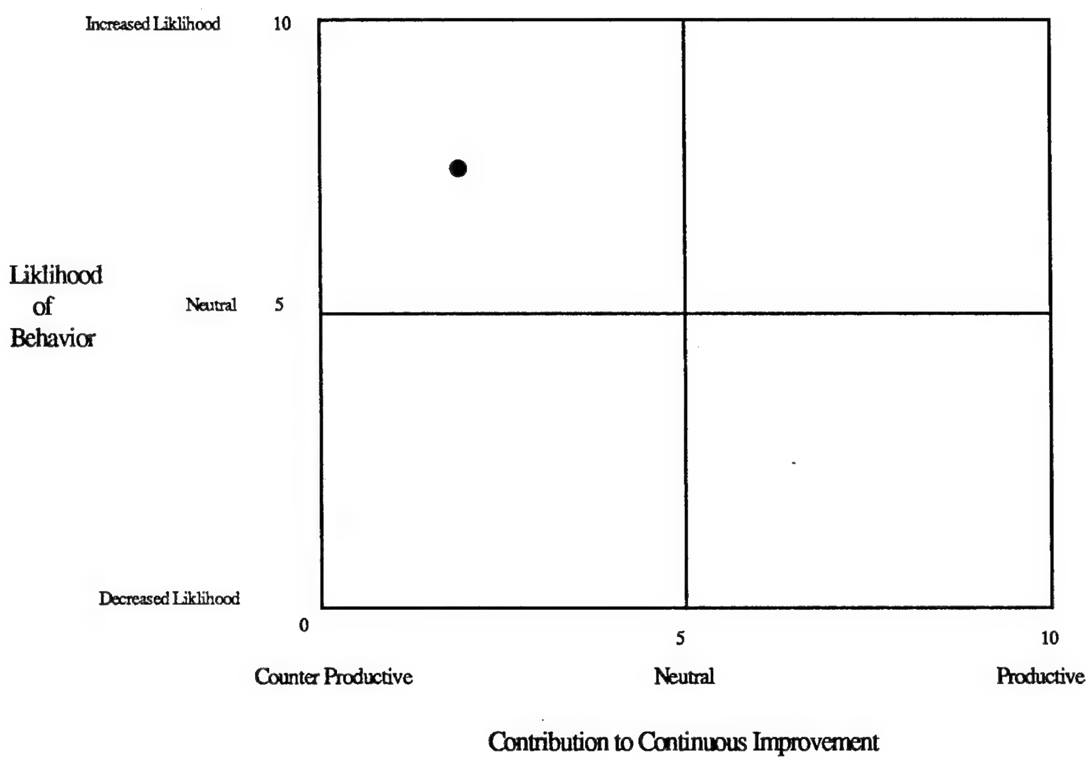

Figure A-6. Metric One, Behavior Six

Behavior 7. Drives improved planning/forecasting methods.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	-	2	-	1	-	7.00
How well does the behavior contribute to CI?	-	-	-	-	-	-	-	4	-	-	8.00

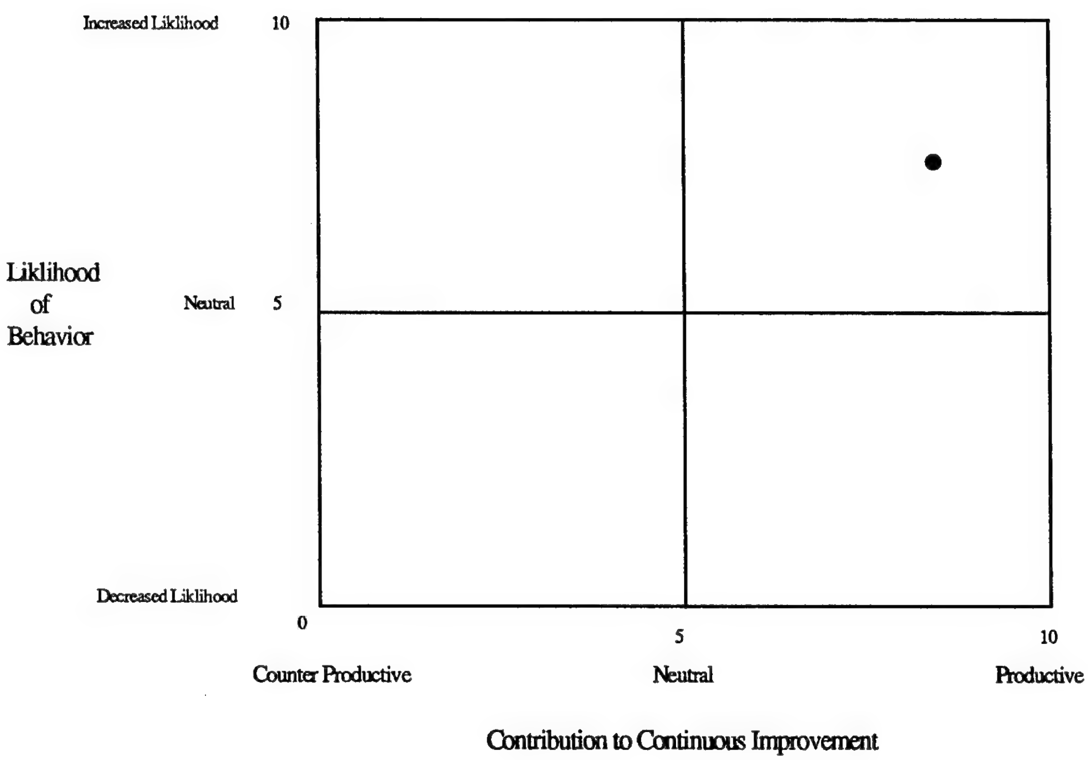

Figure A-7. Metric One, Behavior Seven

Behavior 8. Creates additional work to justify needed management actions.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	-	-	-	1	1	7.00
How well does the behavior contribute to CI?	-	1	-	-	-	-	1	-	1	1	2.00

(Columns 1–10: Participant Ratings)

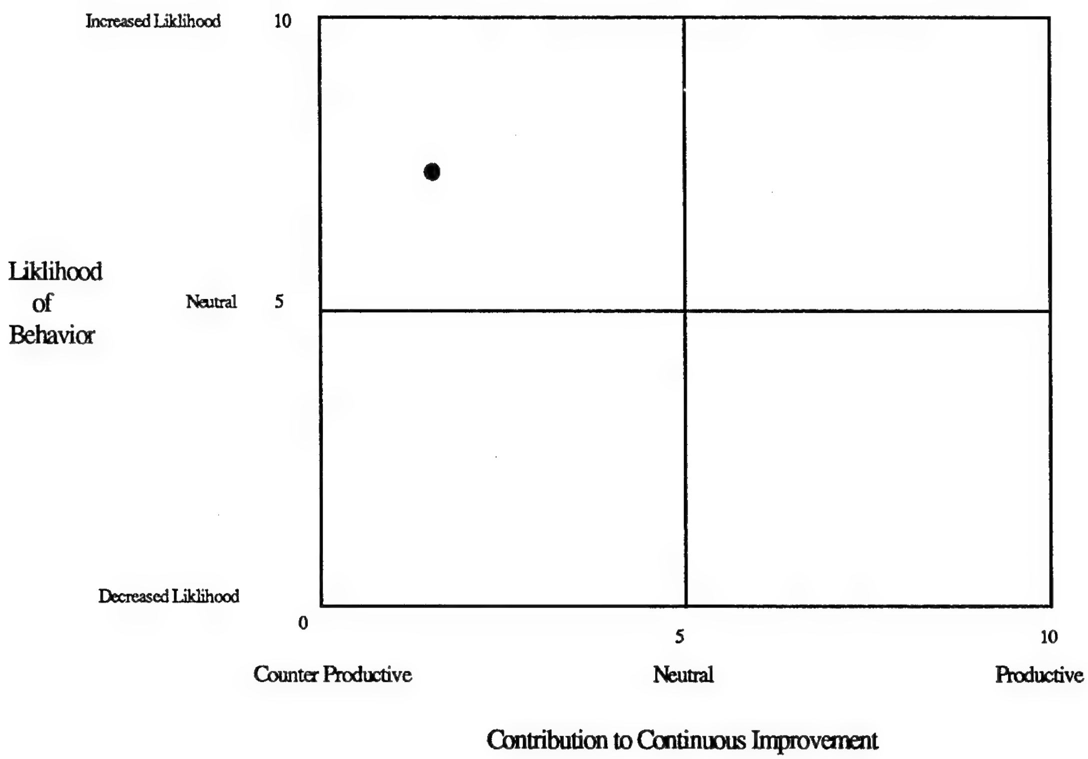

Figure A-8. Metric One, Behavior Eight

Behavior 9. Causes the SPO to plan for technical definition of contracts.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	-	2	-	2	6.75
How well does the behavior contribute to CI?	2	1	1	-	-	-	-	-	-	-	9.00

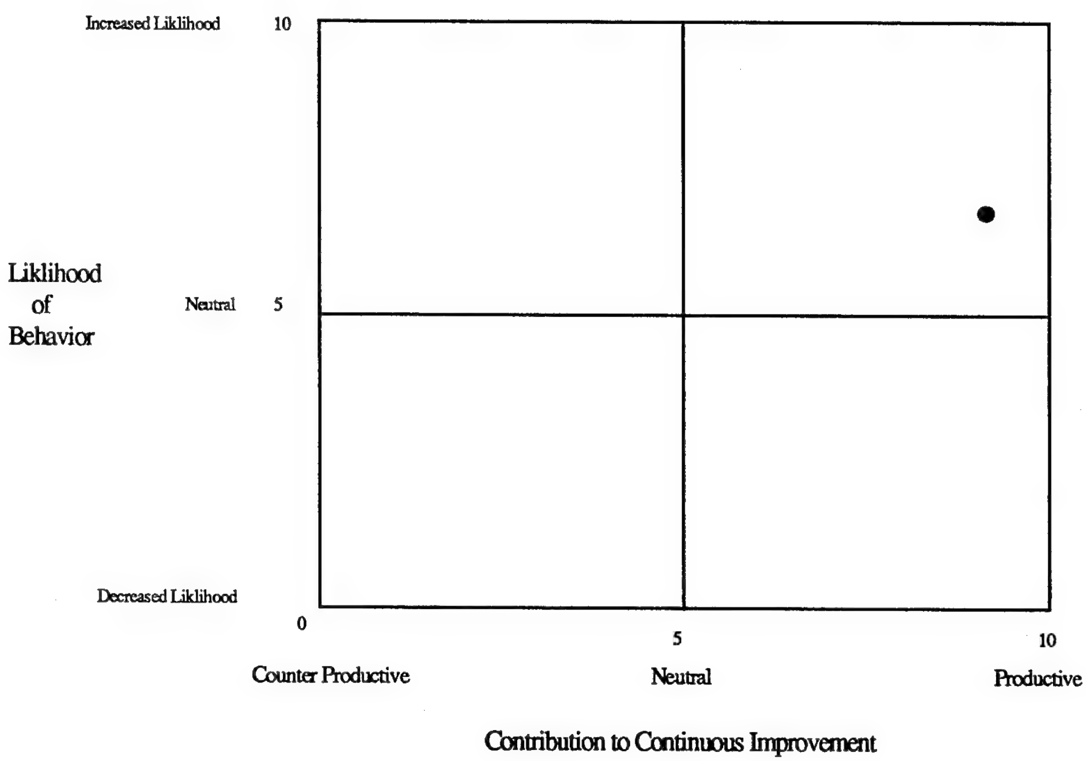

Figure A-9. Metric One, Behavior Nine

Behavior 10. Perform detailed contractor schedule analysis.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	1	-	2	-	-	6.75
How well does the behavior contribute to CI?	-	-	-	-	-	1	3	-	-	-	6.75

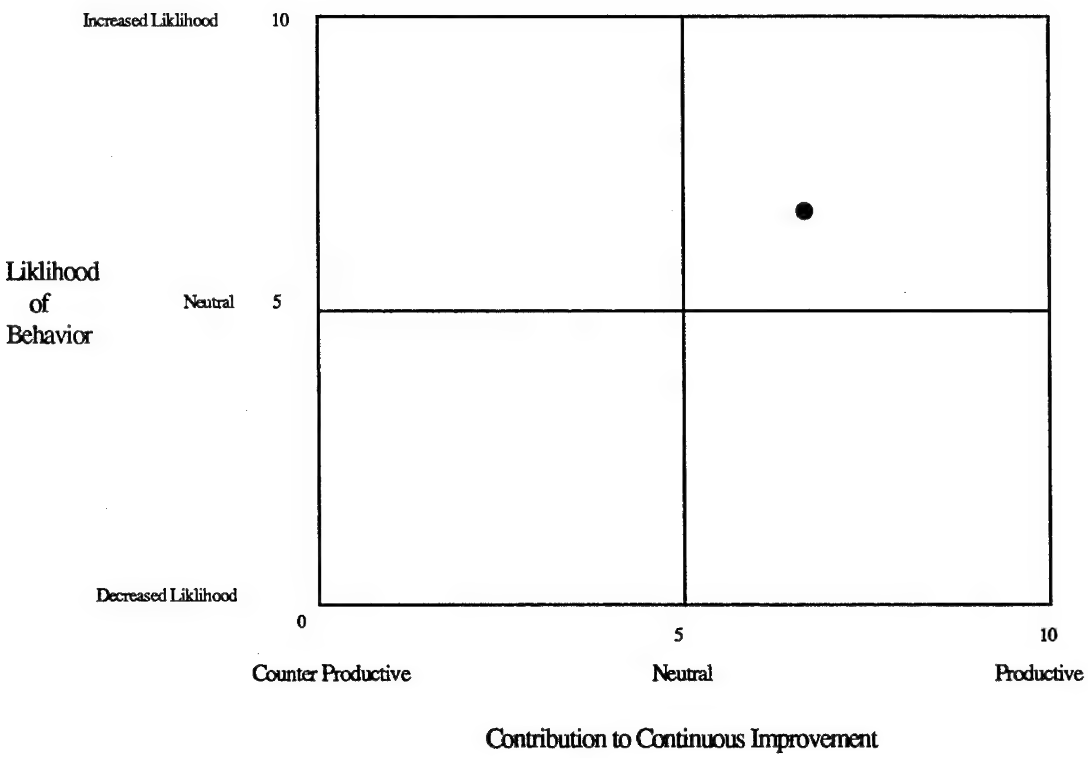

Figure A-10. Metric One, Behavior Ten

Behavior 11. Makes it harder for PMs to handle unforeseen problems.

** Criteria**	Participant Ratings 1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	1	1	-	1	-	6.75
How well does the behavior contribute to CI?	1	1	2	-	-	-	-	-	-	-	2.25

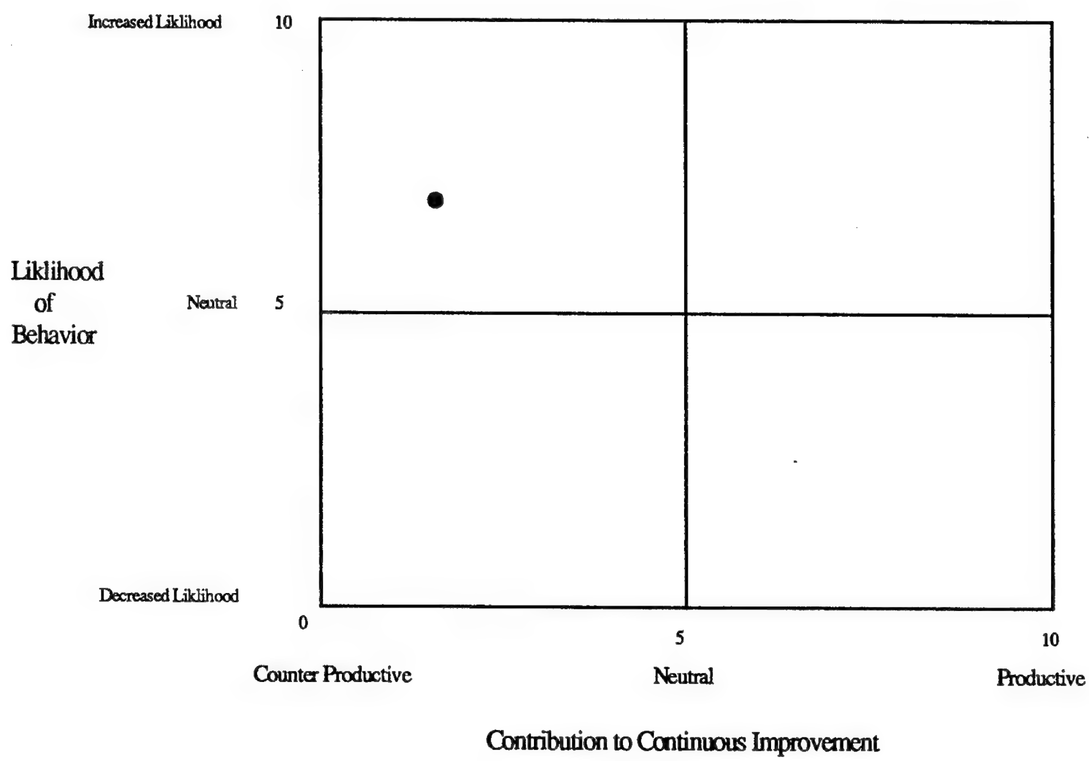

Figure A-11. Metric One, Behavior Eleven

Behavior 12. May cause obligation of funds based on schedule of the metric.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	1	-	-	-	1	1	1	-	6.75
How well does the behavior contribute to CI?	1	1	1	1	-	-	-	-	-	-	2.50

Participant Ratings (columns 1–10)

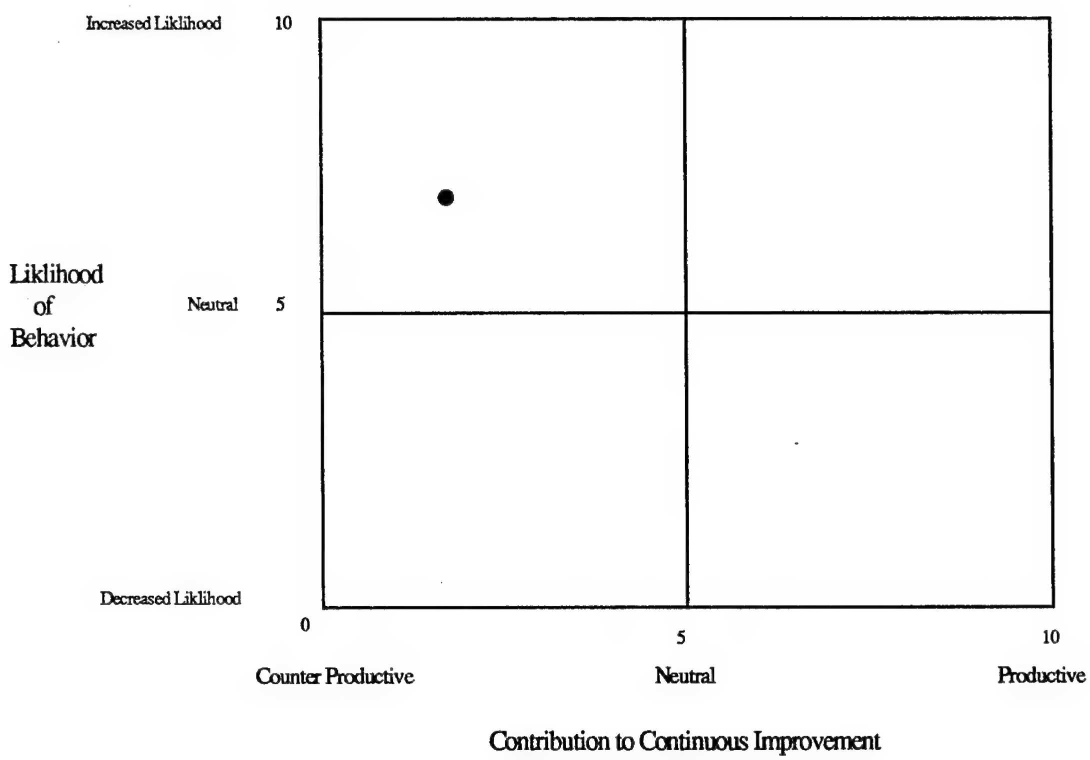

Figure A-12. Metric One, Behavior Twelve

Behavior 13. Offer to return excess funds.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	1	-	1	2	-	-	-	-	5.00
How well does the behavior contribute to CI?	-	-	-	-	-	-	-	1	2	1	9.00

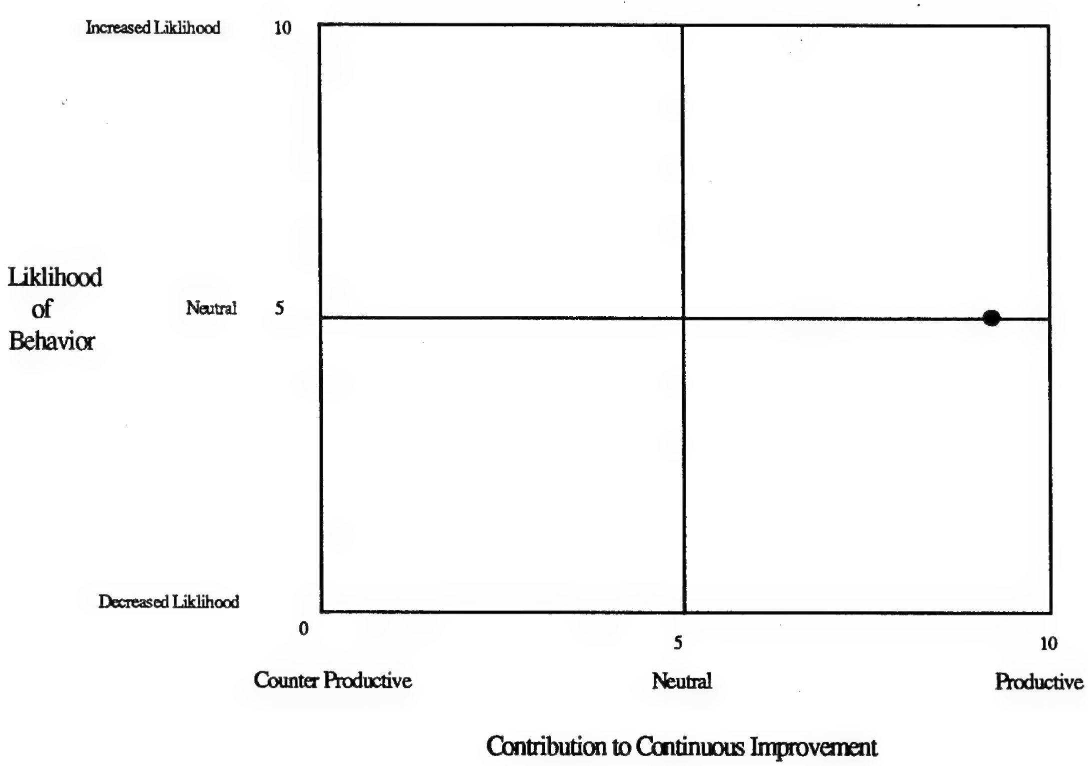

Figure A-13. Metric One, Behavior Thirteen

Behavior 14. Estimates of the planned obligations may be high or low.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	-	-	-	1	1	5.00
How well does the behavior contribute to CI?	1	-	-	-	1	1	-	1	-	-	3.25

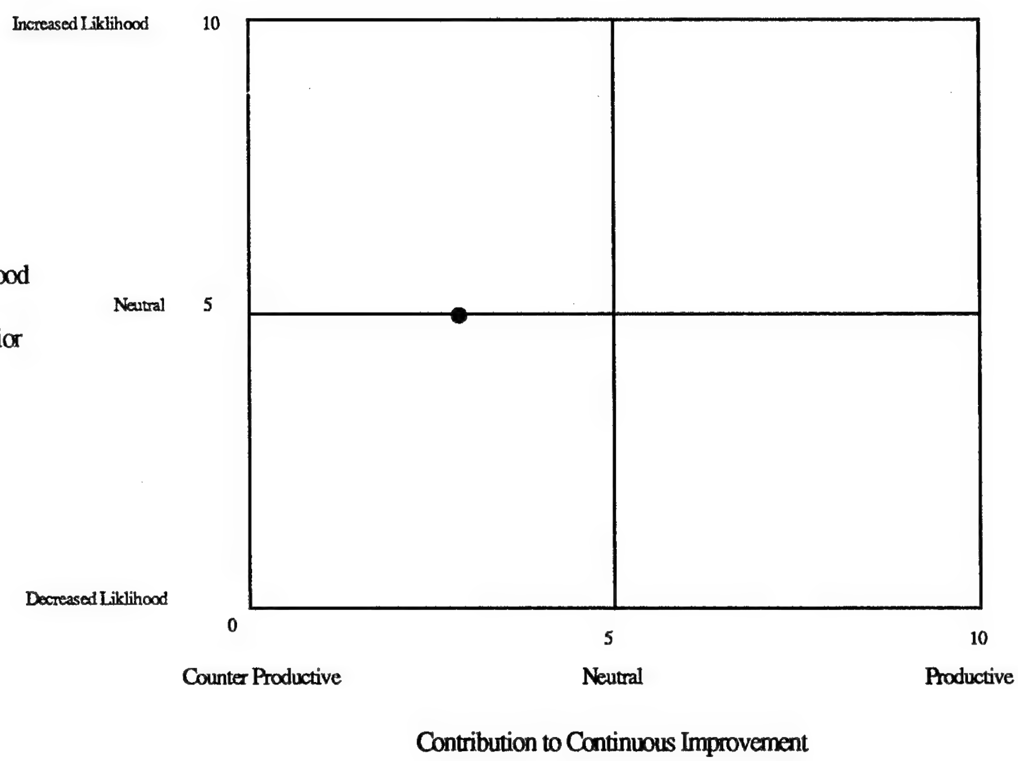

Figure A-14. Metric One, Behavior Fourteen

Metric 2: Required vs. Budgeted Funding Levels

Metric 2 Behavior 1. Brings senior level management attention to programs with big deltas.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	1	1	1	1	8.50
How well does the behavior contribute to CI?	-	-	1	-	1	-	-	2	-	-	6.00

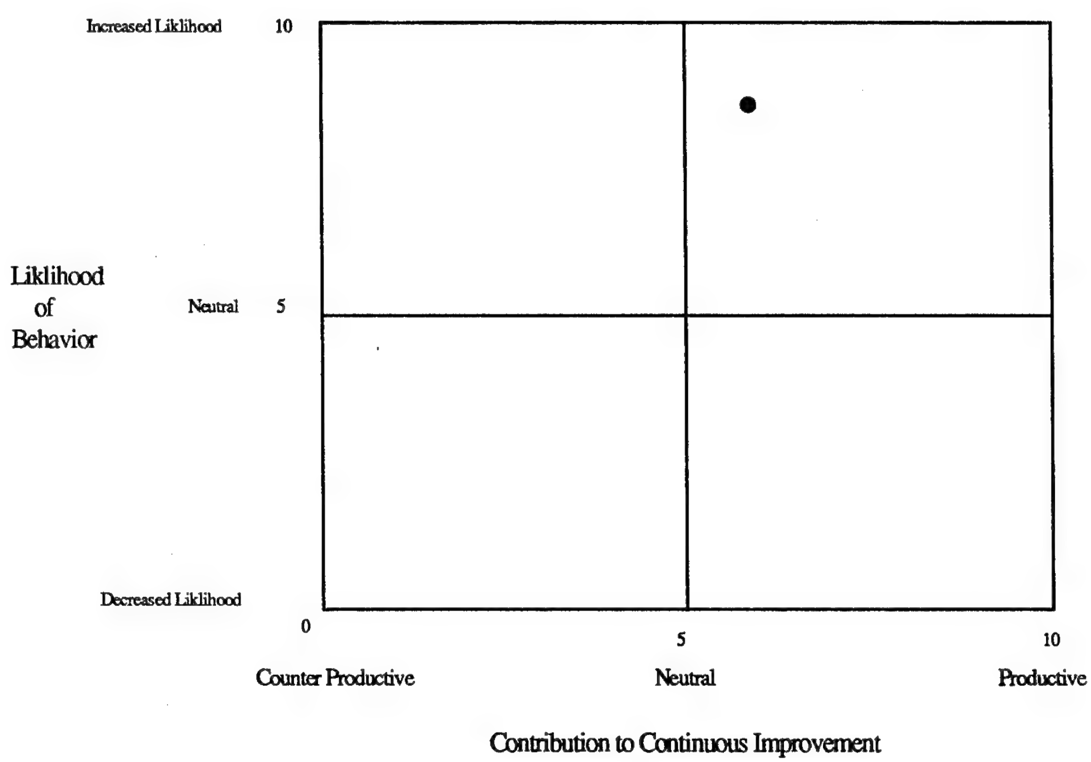

Figure A-15. Metric Two, Behavior One

Behavior 2. Causes identification of excess funds.

** Criteria**

Criteria	Participant Ratings 1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	1	-	1	-	2	8.50
How well does the behavior contribute to CI?	-	-	-	-	-	-	-	2	2	-	8.50

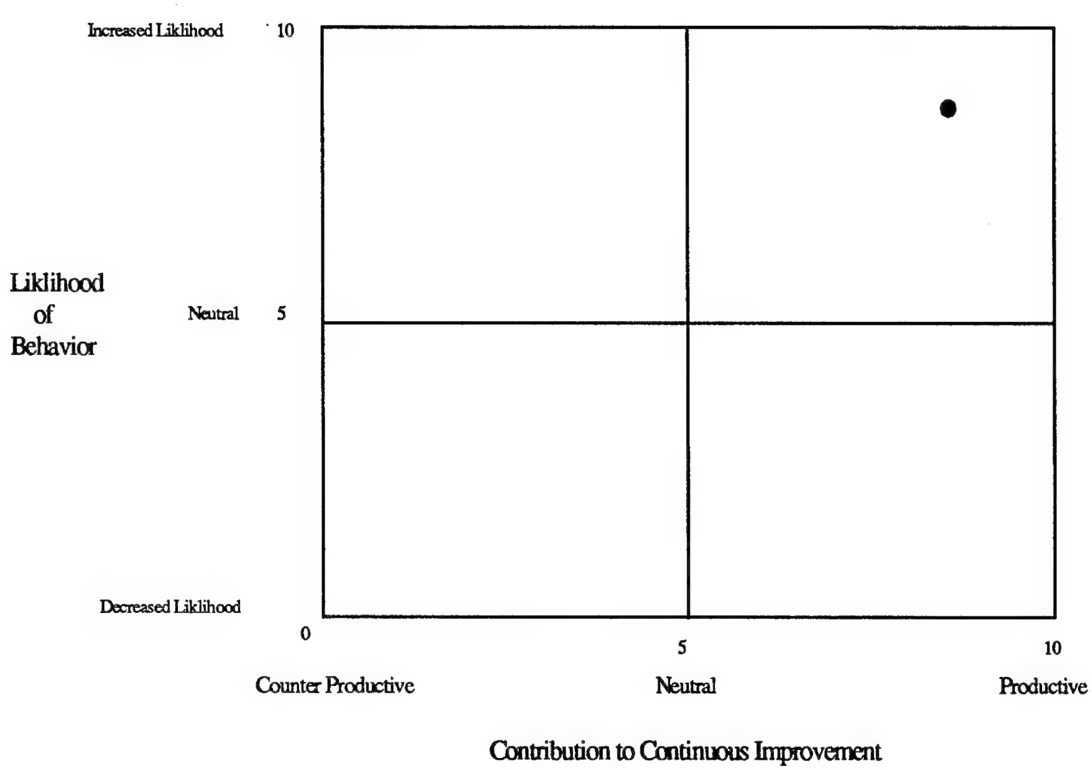

Figure A-16. Metric Two, Behavior Two

Behavior 3. Causes processing of unfunded requirement request.

** Criteria**	Participant Ratings 1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	2	-	1	1	8.25
How well does the behavior contribute to CI?	-	-	1	1	-	-	1	-	-	1	6.00

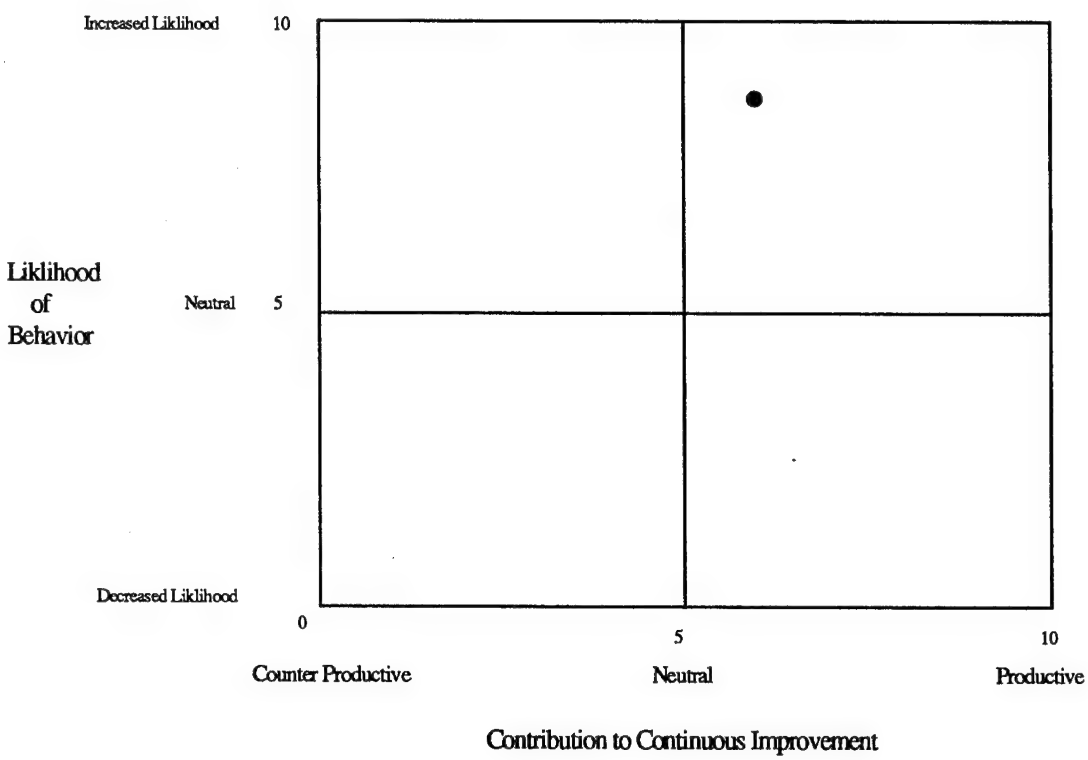

Figure A-17. Metric Two, Behavior Three

Behavior 4. Causes program restructuring.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	1	2	1	-	8.00
How well does the behavior contribute to CI?	1	-	2	-	-	1	-	-	-	-	3.25

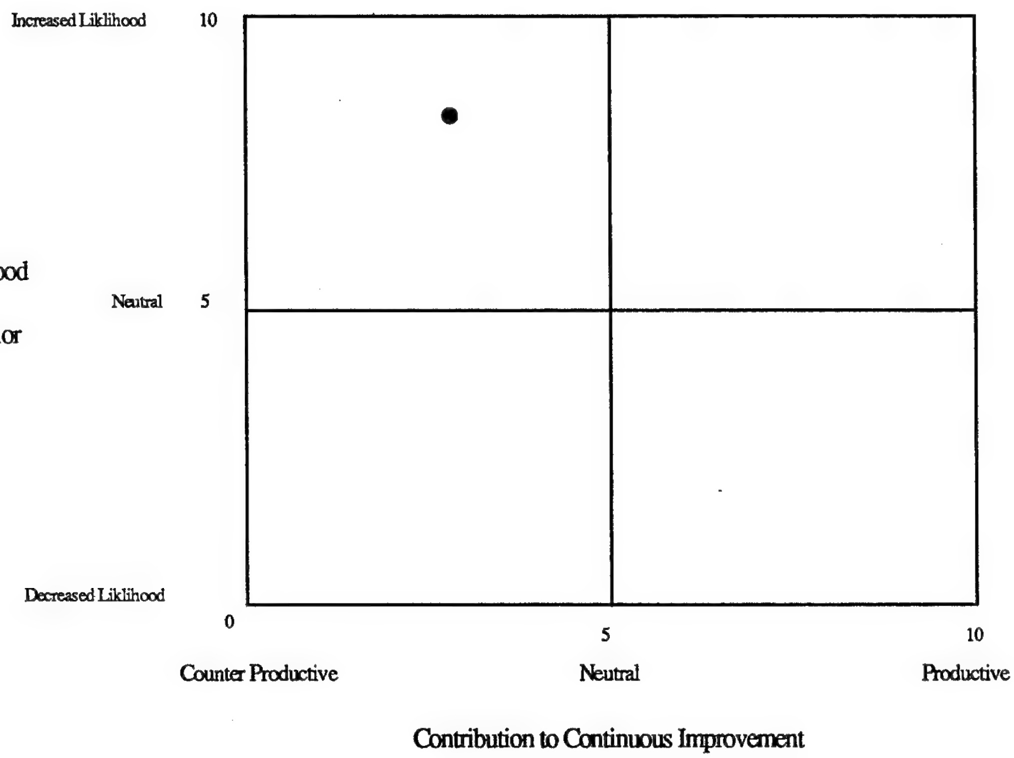

Figure A-18. Metric Two, Behavior Four

Behavior 5. Causes request for congressional reprogramming.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	1	2	1	-	8.00
How well does the behavior contribute to CI?	-	-	2	-	-	1	-	1	-	-	5.00

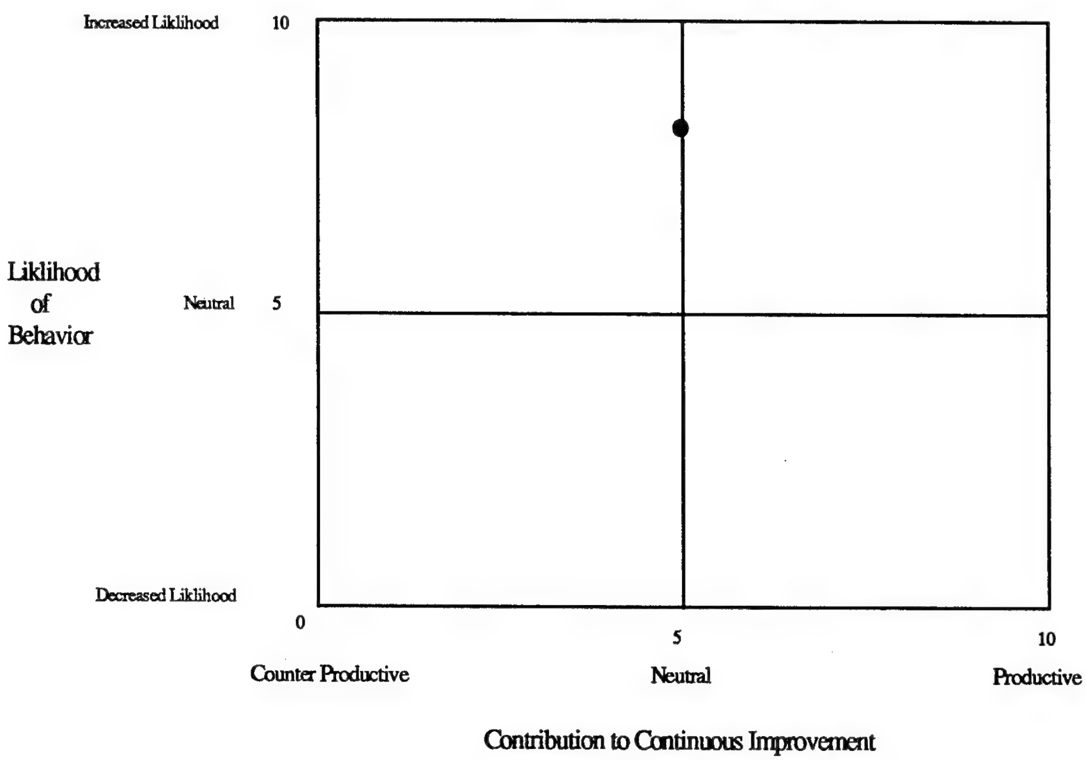

Figure A-19. Metric Two, Behavior Five

Behavior 6. Causes strengthening of budget justification documents.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	1	-	2	1	-	7.75
How well does the behavior contribute to CI?	-	-	-	-	-	-	1	1	1	1	8.50

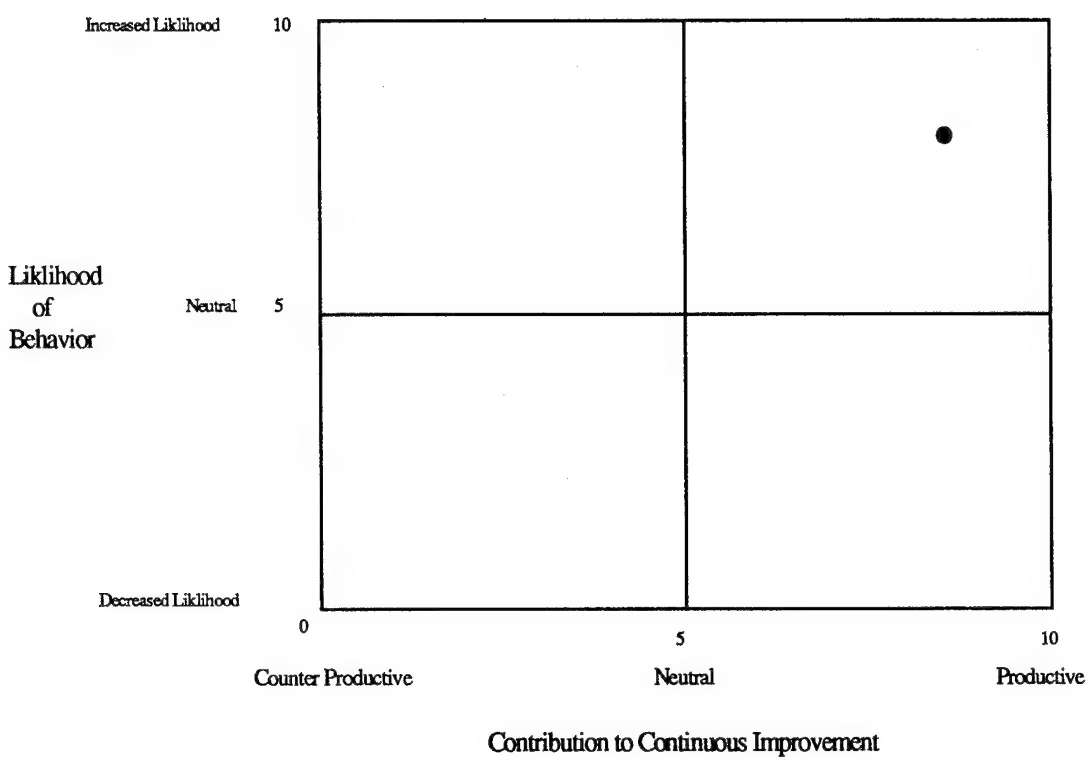

Figure A-20. Metric Two, Behavior Six

Behavior 7. Can cause redistribution of funds.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	1	1	1	-	1	7.75
How well does the behavior contribute to CI?	-	1	-	-	1	1	-	-	1	-	5.50

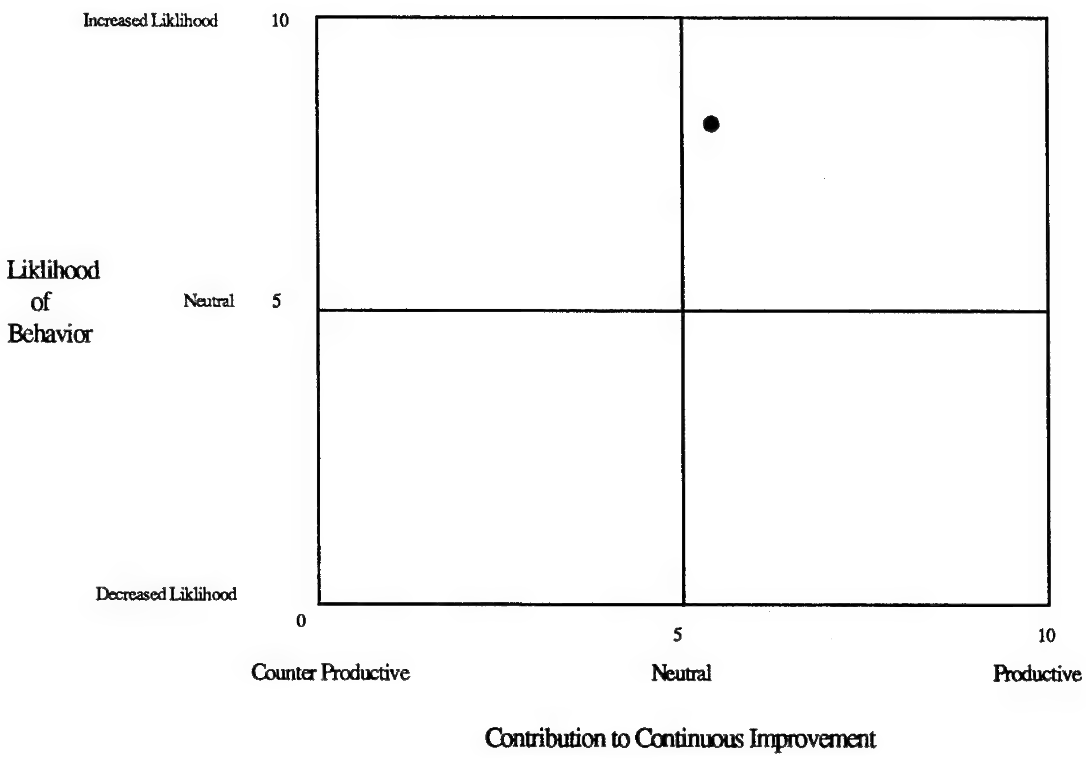

Figure A-21. Metric Two, Behavior Seven

Behavior 8. Cause the SPO to overstate future years requirements.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	-	1	1	1	-	7.25
How well does the behavior contribute to CI?	1	2	-	1	-	-	-	-	-	-	2.25

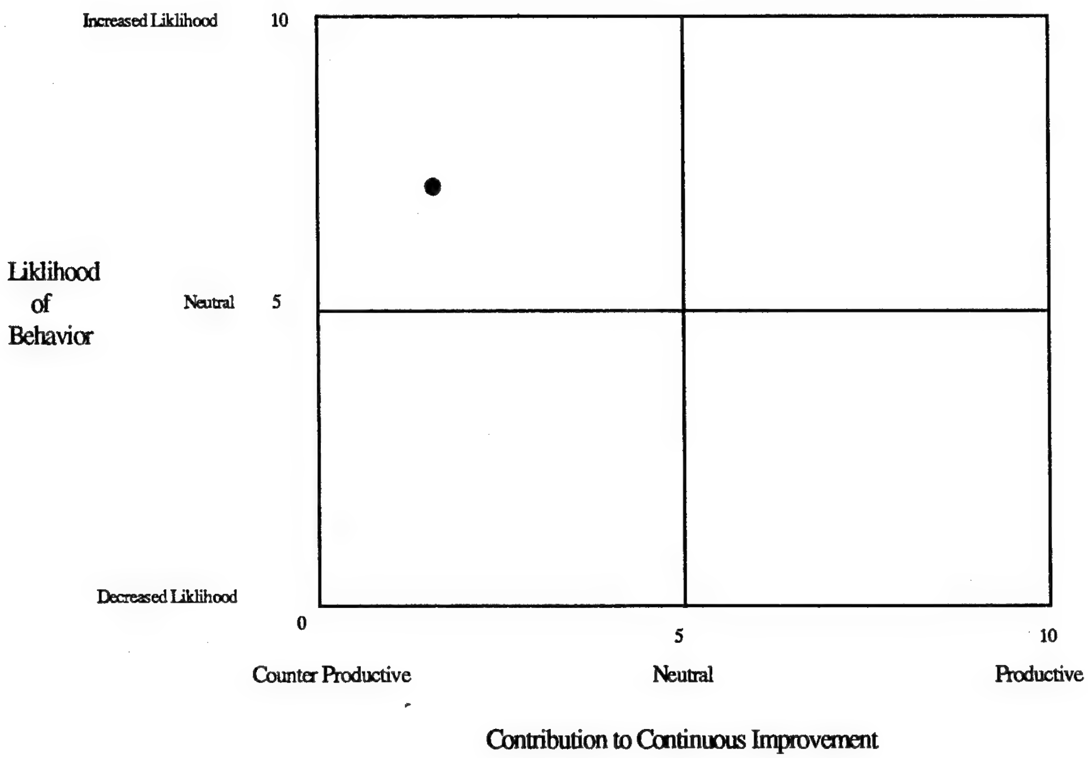

Figure A-22. Metric Two, Behavior Eight

Behavior 9. Causes SPO to create contingency list for "fall-out" funds.

** Criteria**

	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	1	1	-	1	-	6.75
How well does the behavior contribute to CI?	-	-	-	1	-	-	2	-	1	-	6.75

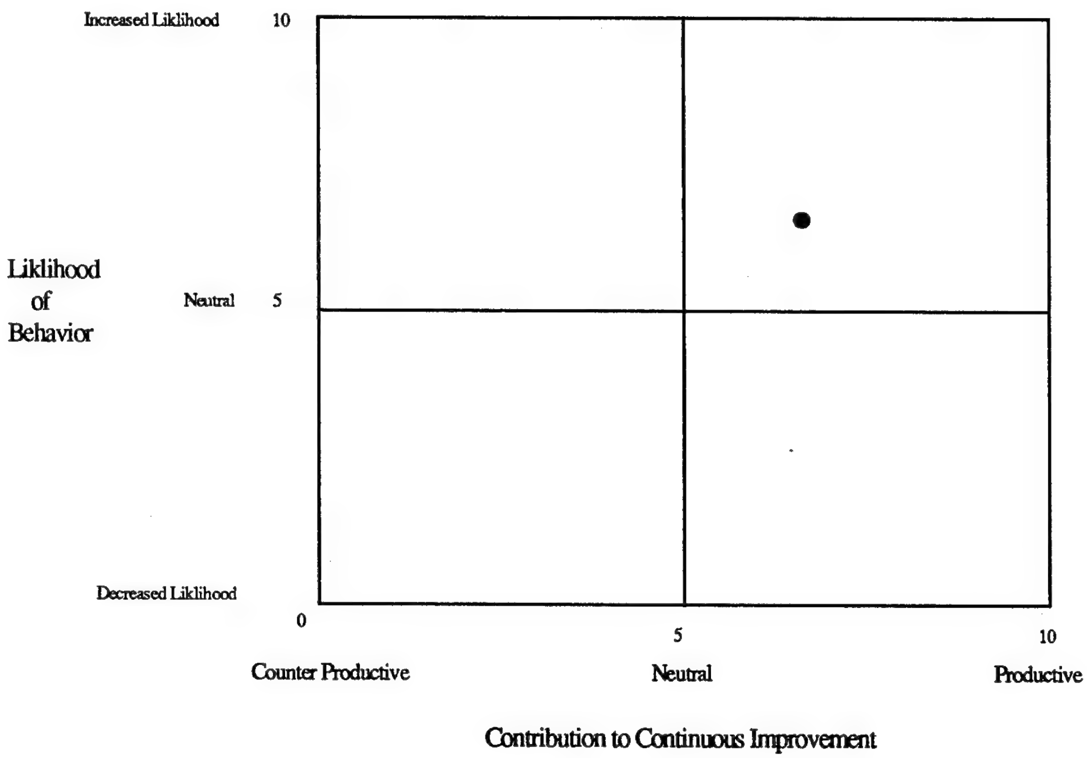

Figure A-23. Metric Two, Behavior Nine

<u>Behavior 10</u>. Causes the SPO to lobby harder for budget.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	1	-	-	1	-	1	1	-	6.50
How well does the behavior contribute to CI?	-	-	-	-	2	-	1	-	1	-	6.50

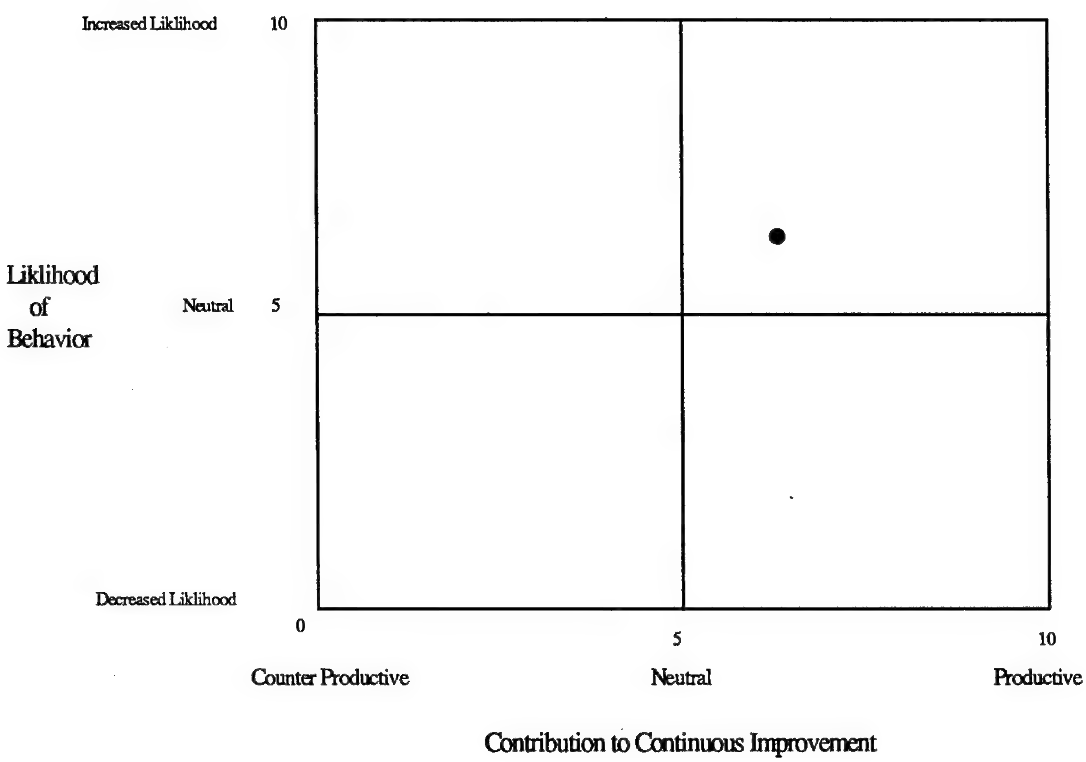

Figure A-24. Metric Two, Behavior Ten

Behavior 11. Causes partial or total contract termination.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	1	-	2	1	-	-	-	5.75
How well does the behavior contribute to CI?	1	1	1	-	1	-	-	-	-	-	2.75

(Columns 1–10: Participant Ratings)

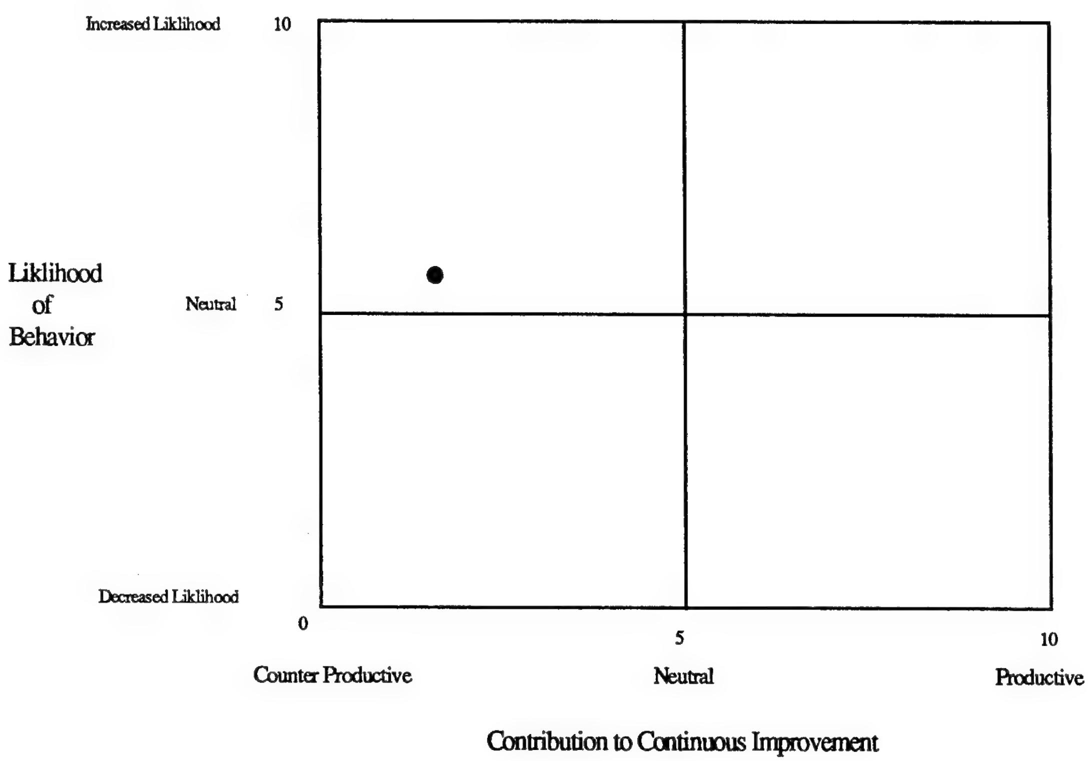

Figure A-25. Metric Two, Behavior Eleven

Behavior 12. Causes managers to violate appropriation law.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	4	-	-	-	-	-	5.00
How well does the behavior contribute to CI?	4	-	-	-	-	-	-	-	-	-	1.00

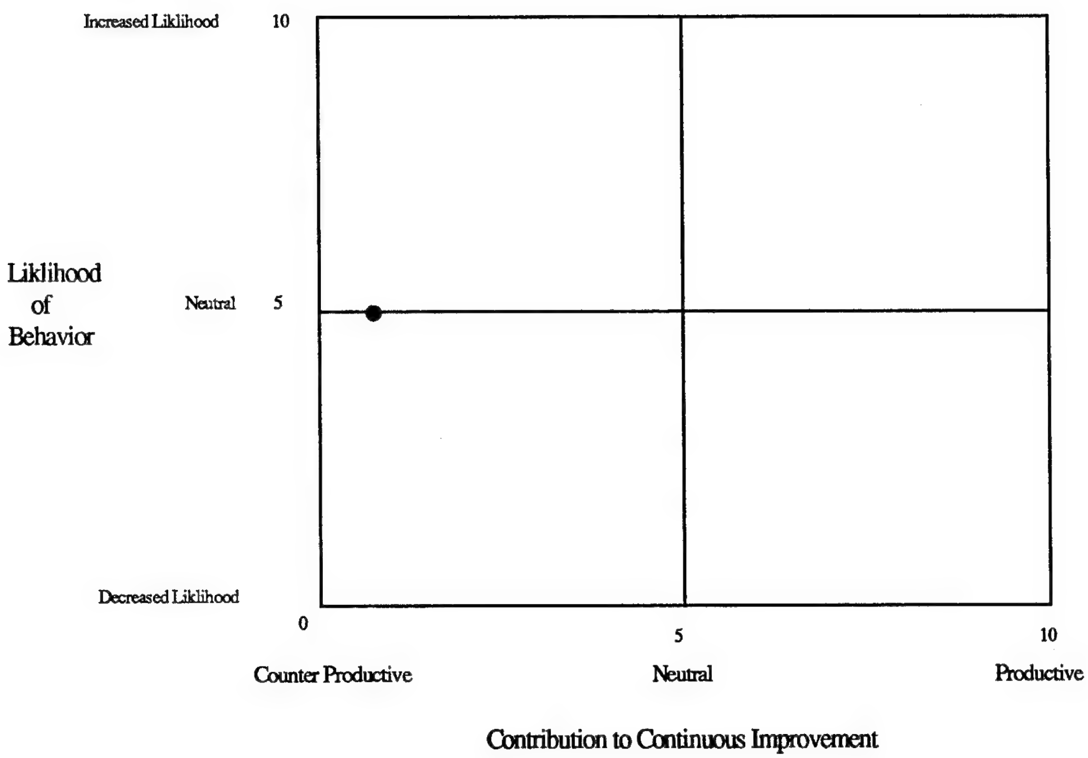

Figure A-26. Metric Two, Behavior Twelve

Behavior 13. Causes procurement of unauthorized goods.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
	Participant Ratings										
How well does the metric drive the behavior?	-	-	1	-	2	1	-	-	-	-	4.75
How well does the behavior contribute to CI?	3	1	-	-	-	-	-	-	-	-	1.25

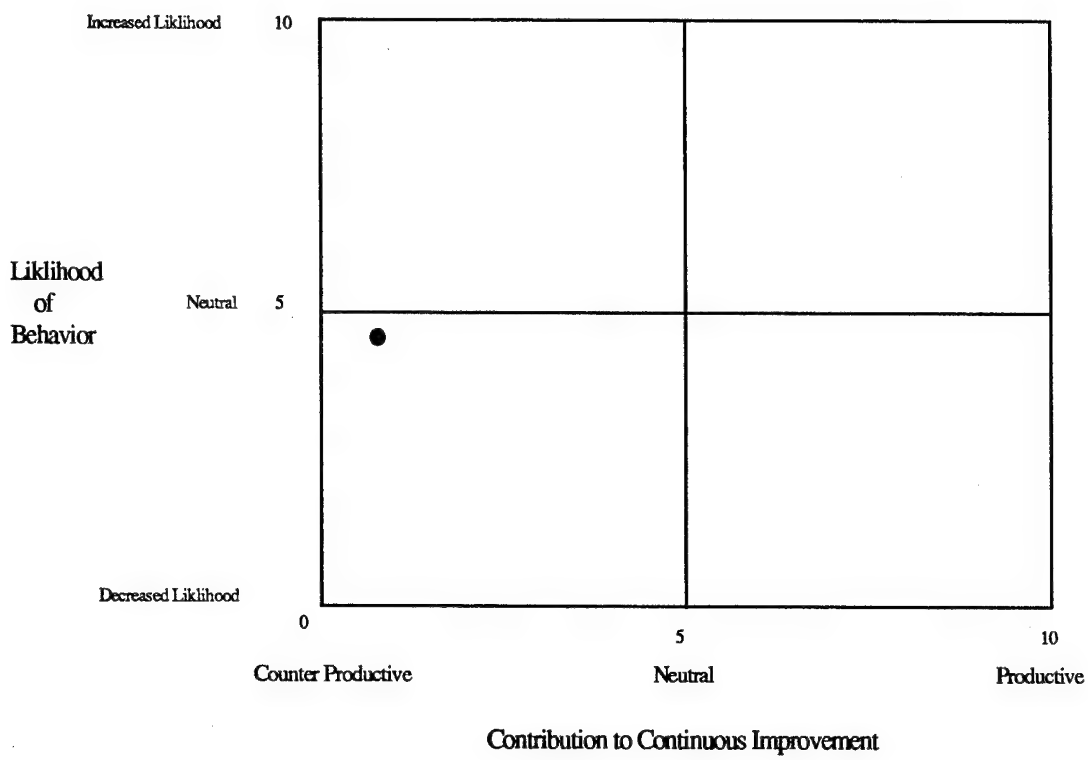

Figure A-27. Metric Two, Behavior Thirteen

Metric 3: Percent Cost Variance

Metric 3 Behavior 1. Causes the contractor to explain the variance.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	-	1	1	2	9.25
How well does the behavior contribute to CI?	-	-	-	1	-	1	1	1	-	-	6.25

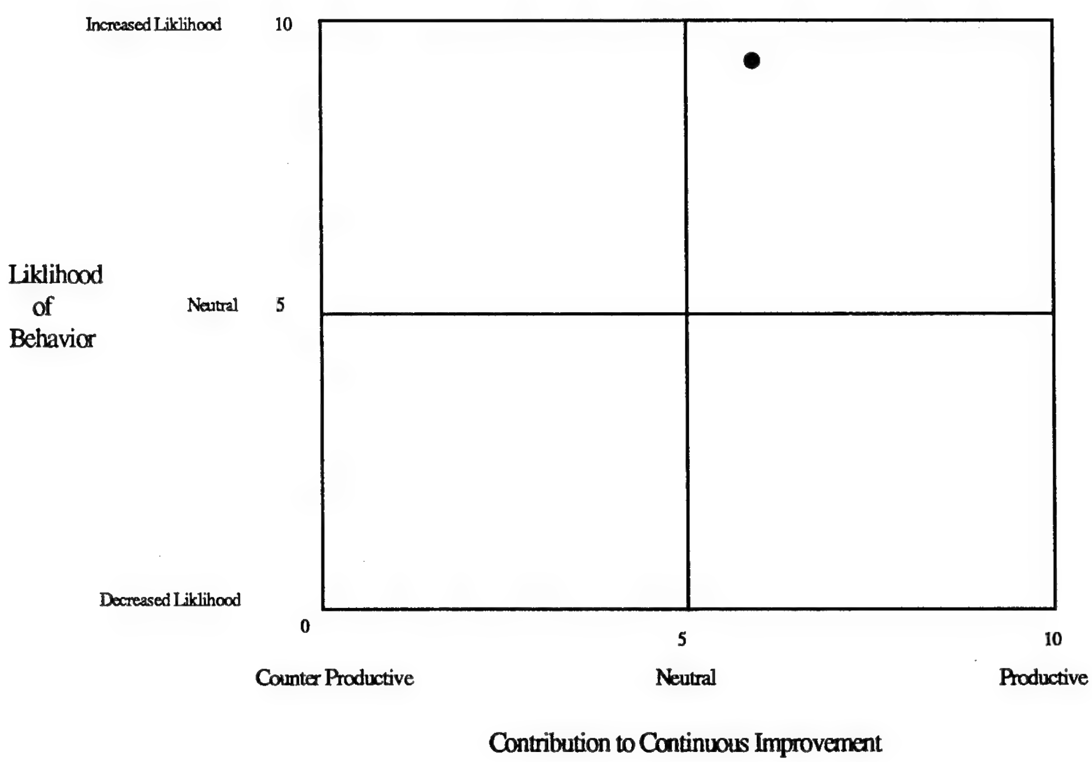

Figure A-28. Metric Three, Behavior One

Behavior 2. Creates additional work to justify deviations.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	-	1	1	2	9.25
How well does the behavior contribute to CI?	-	-	2	1	-	-	1	-	-	-	4.25

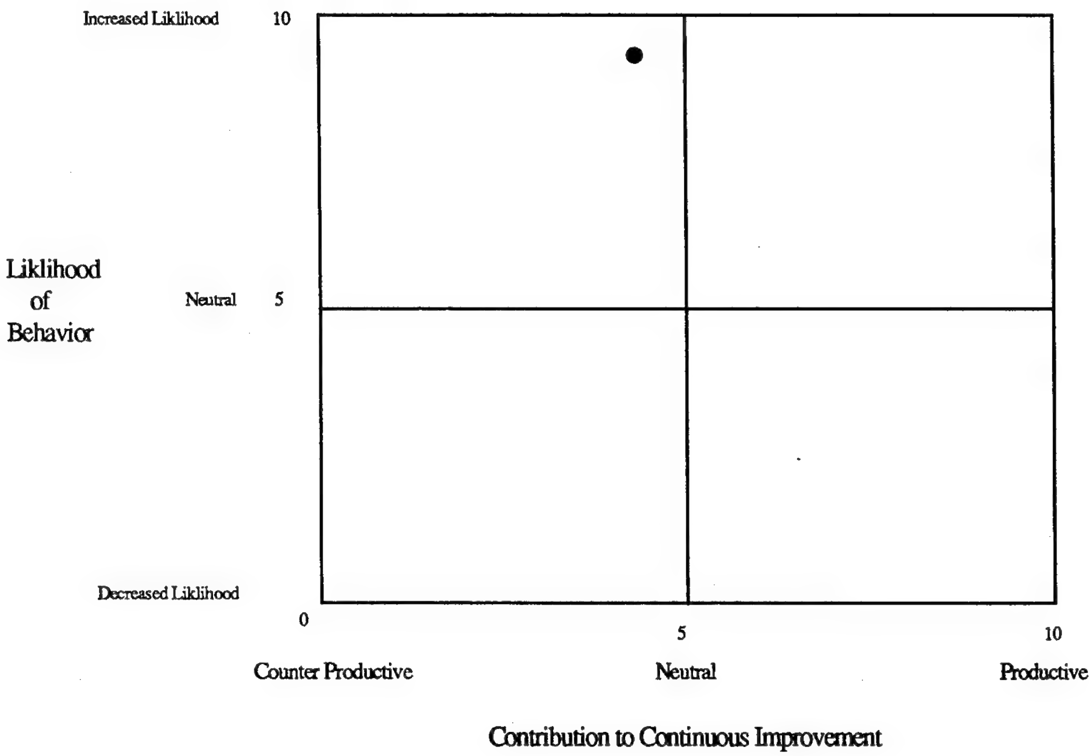

Figure A-29. Metric Three, Behavior Two

Behavior 3. Causes contractor to apply management reserve to overruns.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
	Participant Ratings										
How well does the metric drive the behavior?	-	-	-	-	-	-	1	-	2	1	8.75
How well does the behavior contribute to CI?	-	-	-	1	1	1	-	1	-	-	5.75

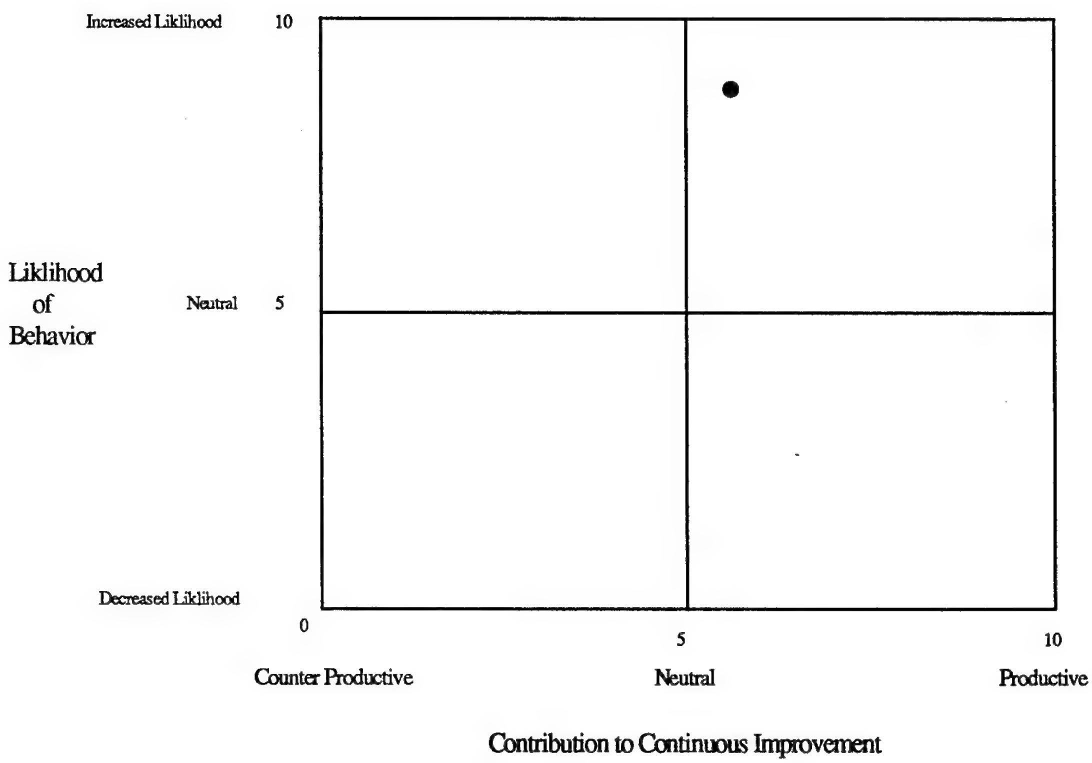

Figure A-30. Metric Three, Behavior Three

Behavior 4. Compare contract overrun with budget status.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	1	-	2	1	8.75
How well does the behavior contribute to CI?	-	-	-	-	2	1	-	1	-	-	6.00

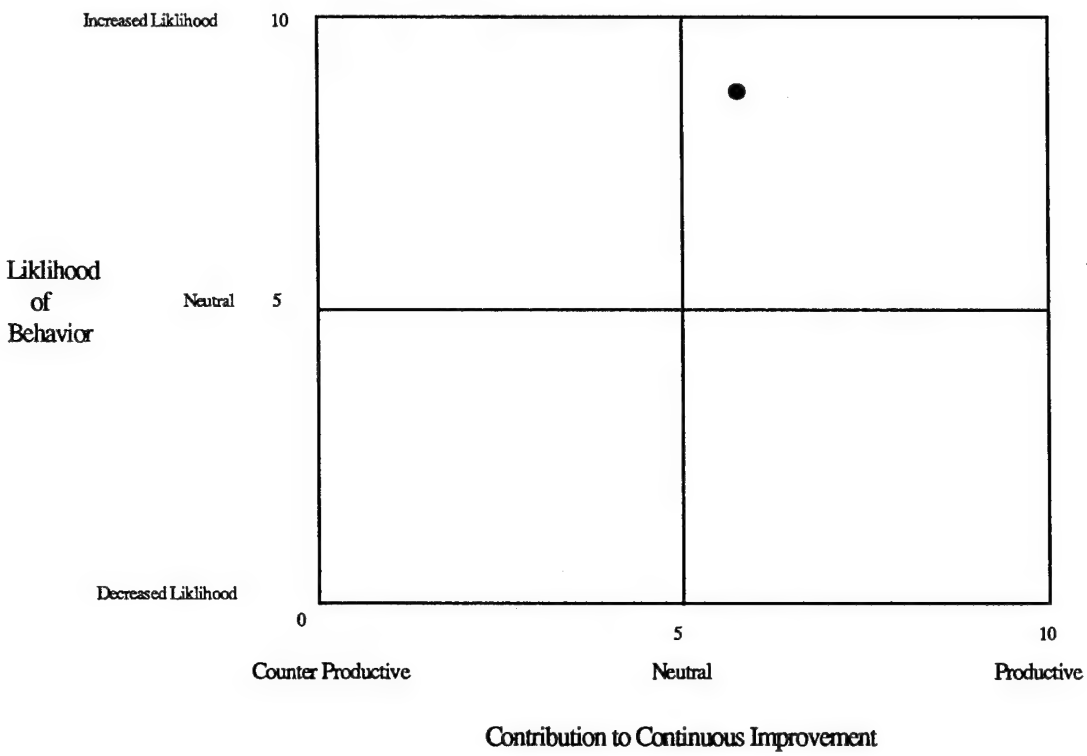

Figure A-31. Metric Three, Behavior Four

Behavior 5. Causes increased oversight of contractors.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	1	-	-	2	1	8.50
How well does the behavior contribute to CI?	-	-	-	3	1	-	-	-	-	-	4.25

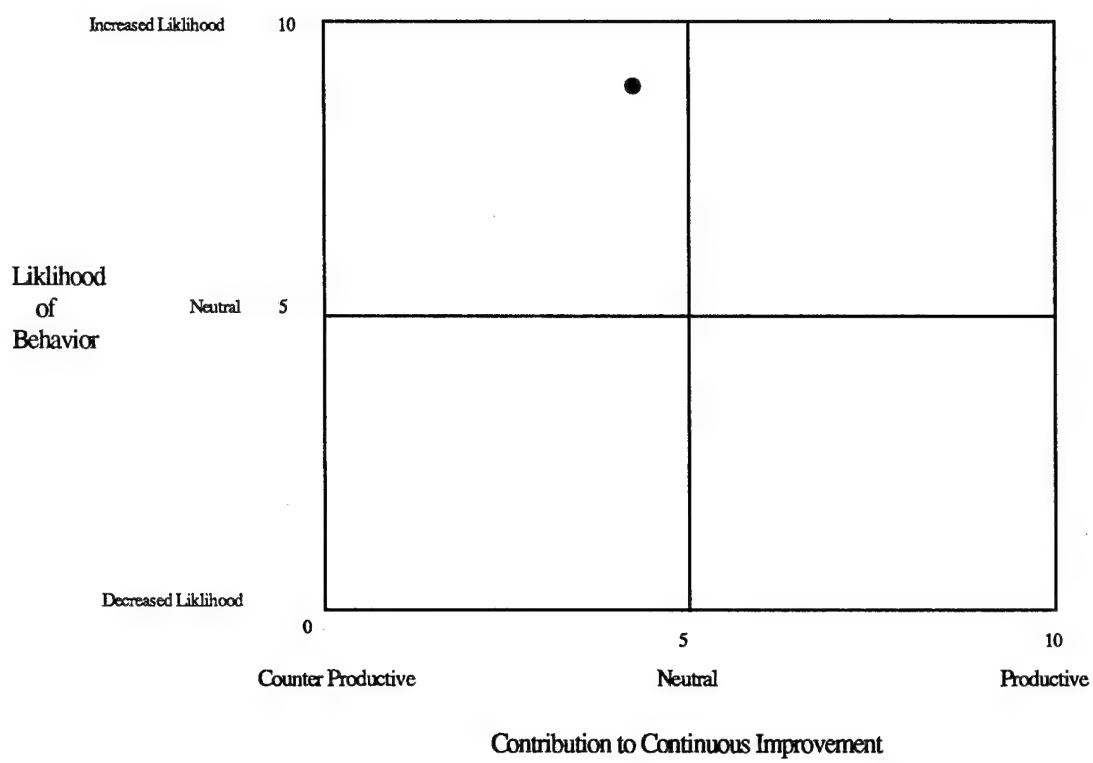

Figure A-32. Metric Three, Behavior Five

Behavior 6. Causes the SPO to look into variances.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	1	-	2	1	8.25
How well does the behavior contribute to CI?	-	-	1	-	-	-	1	1	1	-	6.75

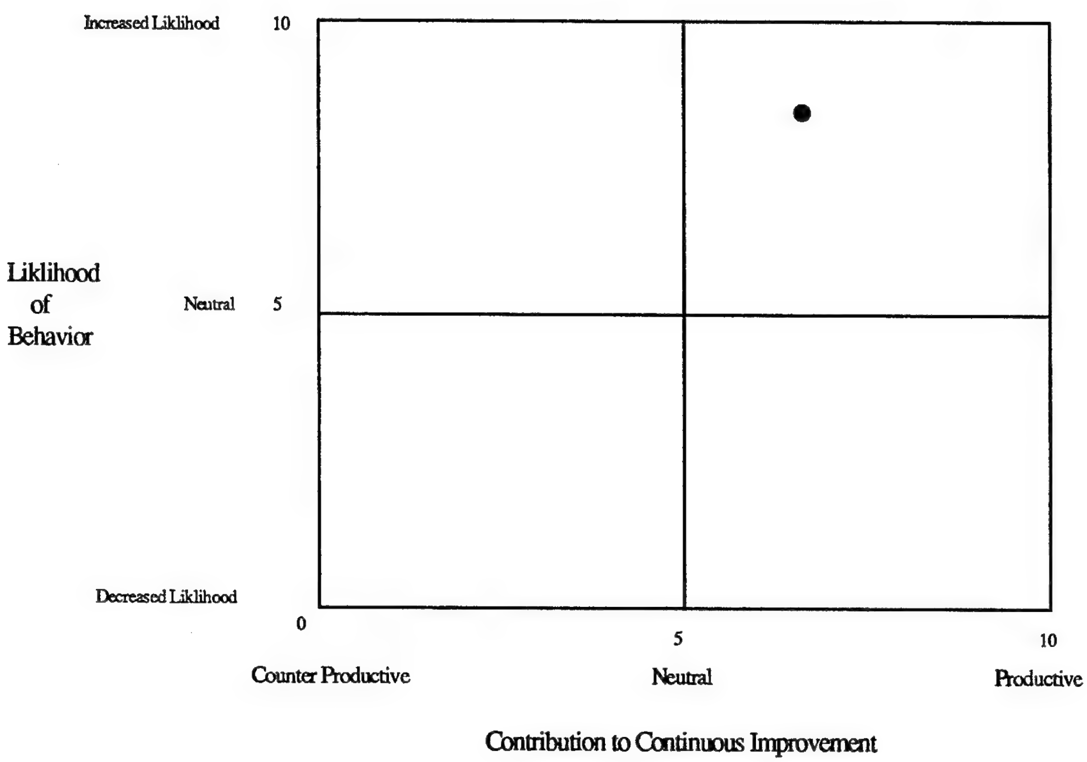

Figure A-33. Metric Three, Behavior Six

Behavior 7. Cost cutting measures.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	1	-	1	1	1	8.25
How well does the behavior contribute to CI?	-	-	-	-	1	2	-	1	-	-	6.25

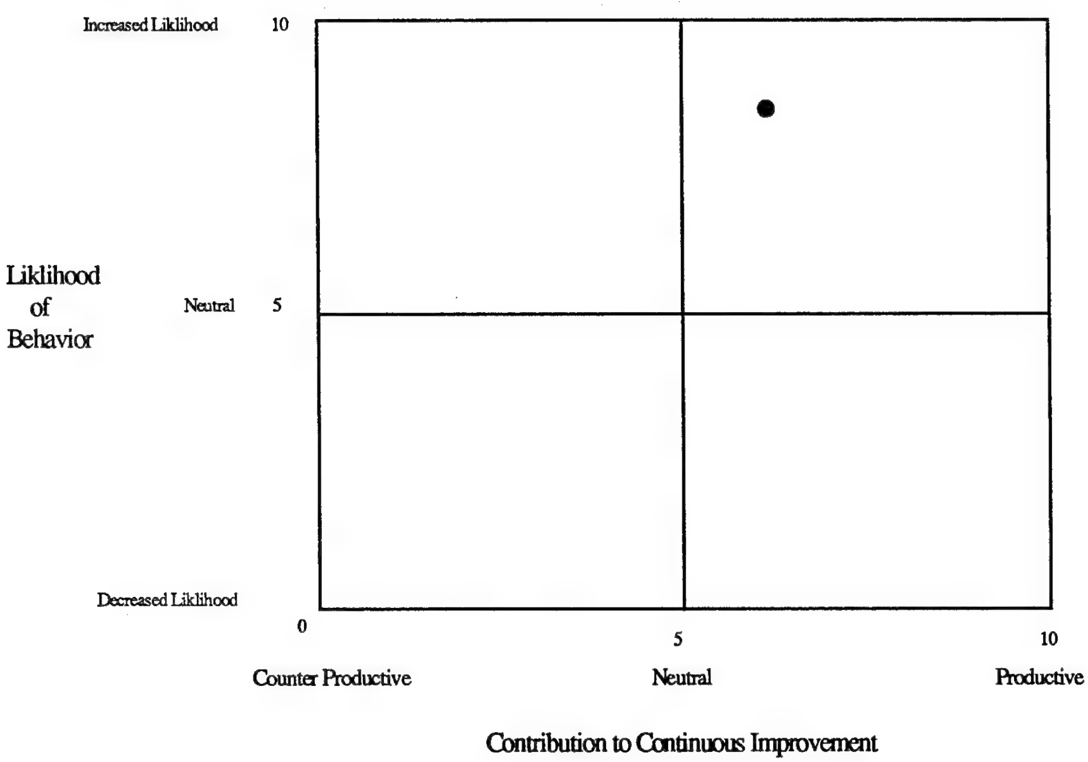

Figure A-34. Metric Three, Behavior Seven

Behavior 8. Budget for anticipated overruns.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	2	1	1	-	7.75
How well does the behavior contribute to CI?	-	-	-	1	-	-	-	3	-	-	7.00

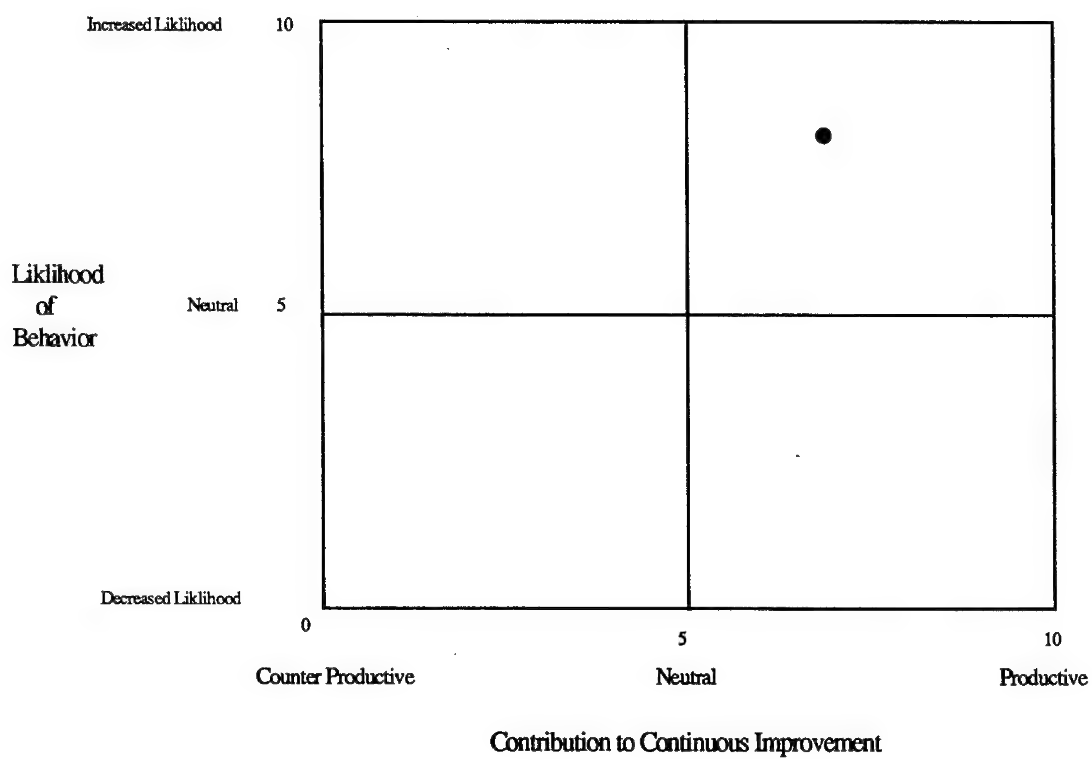

Figure A-35. Metric Three, Behavior Eight

Behavior 9. Helps to control contractor spending.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	1	1	1	1	-	7.50
How well does the behavior contribute to CI?	-	-	-	-	1	-	-	2	-	1	7.75

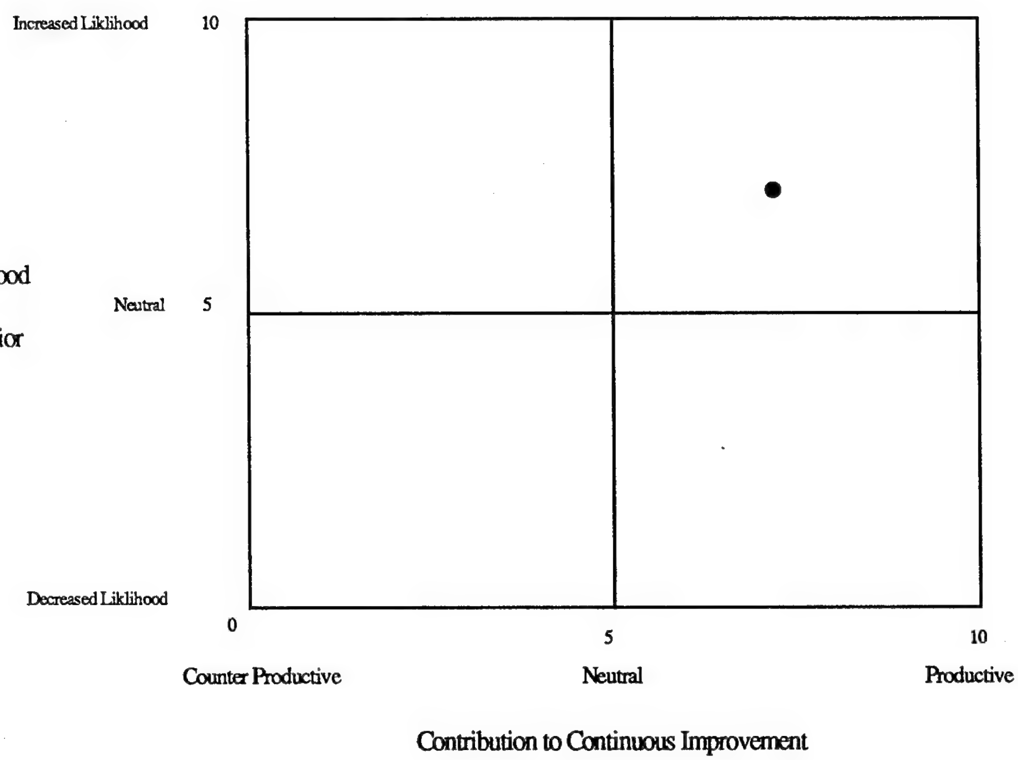

Figure A-36. Metric Three, Behavior Nine

Behavior 10. Causes the SPO/contractor to rebaseline.

** Criteria**	Participant Ratings 1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	2	-	1	1	-	7.25
How well does the behavior contribute to CI?	-	-	2	2	-	-	-	-	-	-	3.50

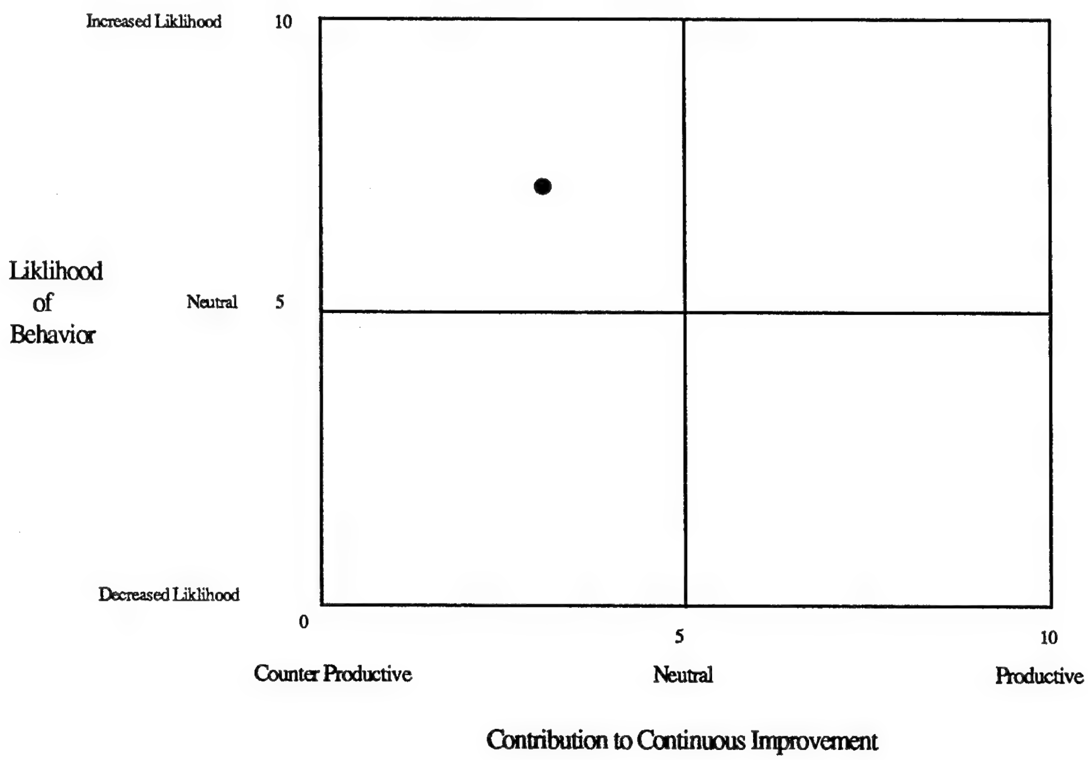

Figure A-37. Metric Three, Behavior Ten

<u>Behavior 11</u>. Helps to keep programs under control.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	-	1	1	1	-	7.25
How well does the behavior contribute to CI?	-	-	-	-	-	1	-	1	-	2	8.50

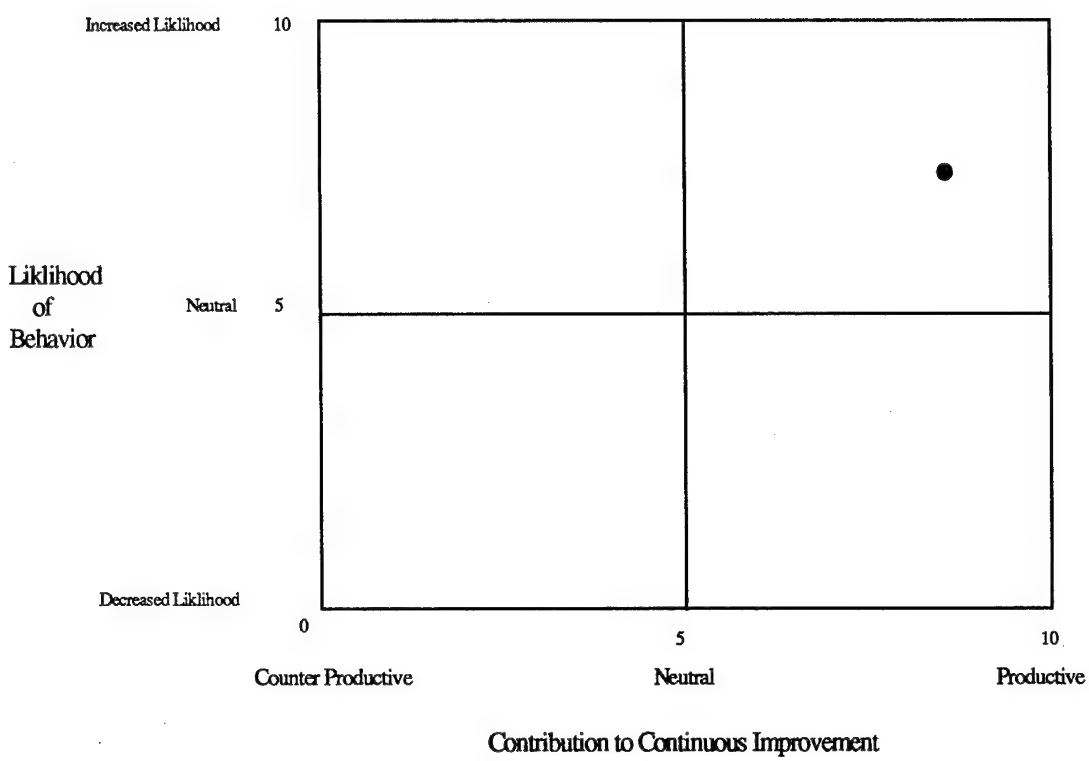

Figure A-38. Metric Three, Behavior Eleven

Behavior 12. Causes examination of schedule performance.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
					Participant Ratings						
How well does the metric drive the behavior?	-	-	-	-	-	1	2	1	-	-	7.00
How well does the behavior contribute to CI?	-	-	-	-	-	-	-	3	1	-	8.25

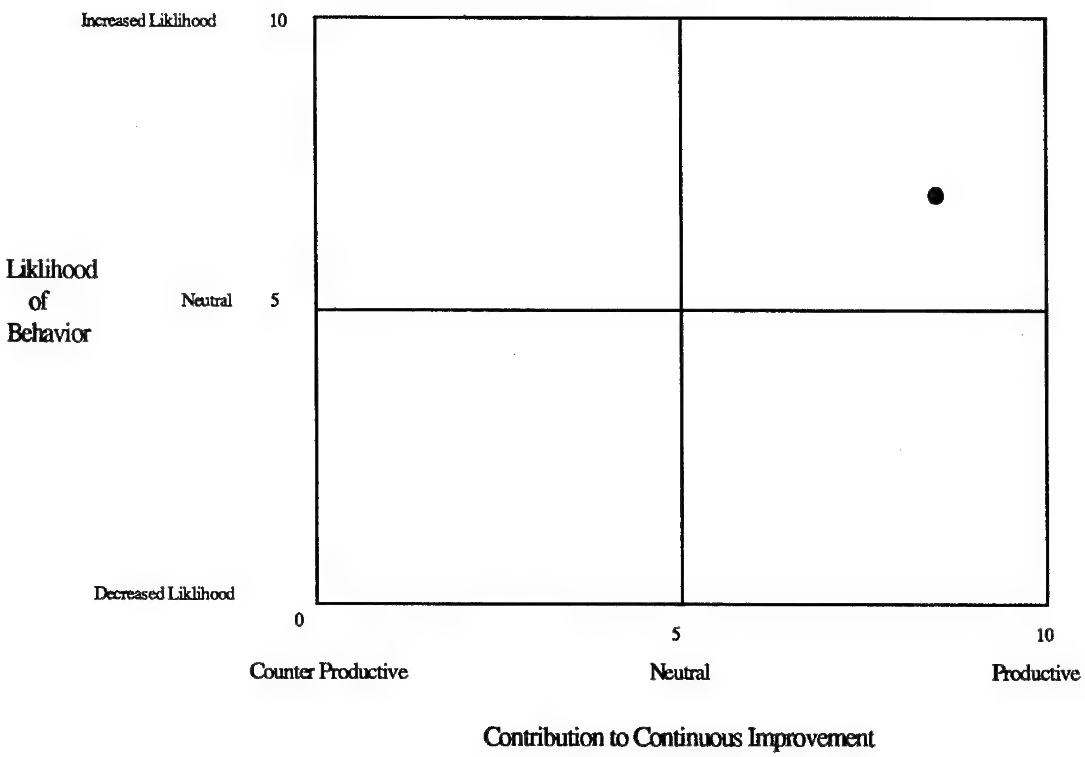

Figure A-39. Metric Three, Behavior Twelve

Behavior 13. Causes the SPO to watch overhead rates.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	2	-	2	-	-	7.00
How well does the behavior contribute to CI?	-	-	-	1	2	1	-	-	-	-	5.00

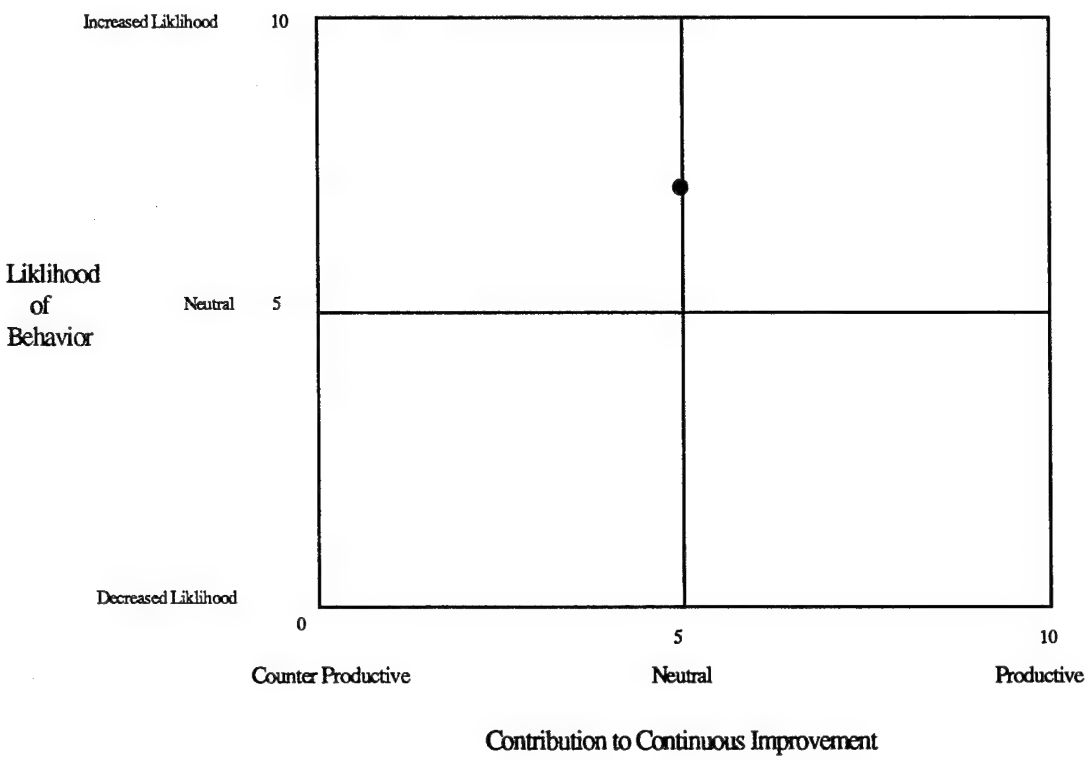

Figure A-40. Metric Three, Behavior Thirteen

Behavior 14. Causes the SPO to mandate cost cutting.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	1	-	1	1	-	7.00
How well does the behavior contribute to CI?	-	-	2	2	-	-	-	-	-	-	3.50

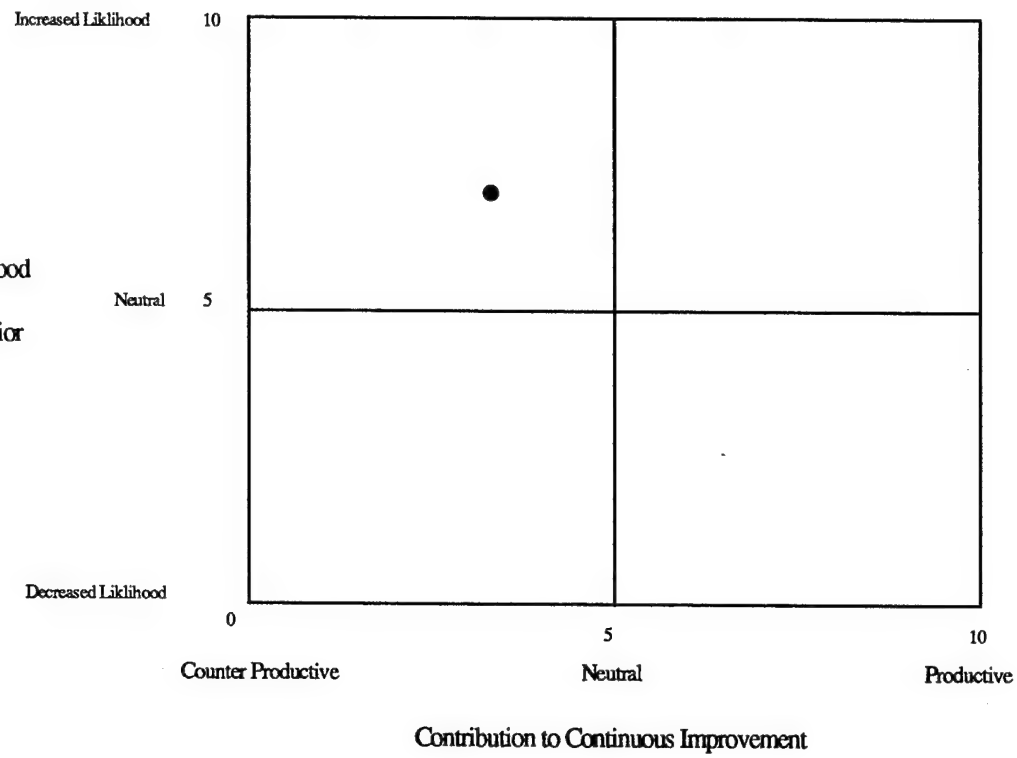

Figure A-41. Metric Three, Behavior Fourteen

Behavior 15. Revise billing, liquidation, and profit rates.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	1	3	-	-	-	6.75
How well does the behavior contribute to CI?	-	-	-	2	1	1	-	-	-	-	4.75

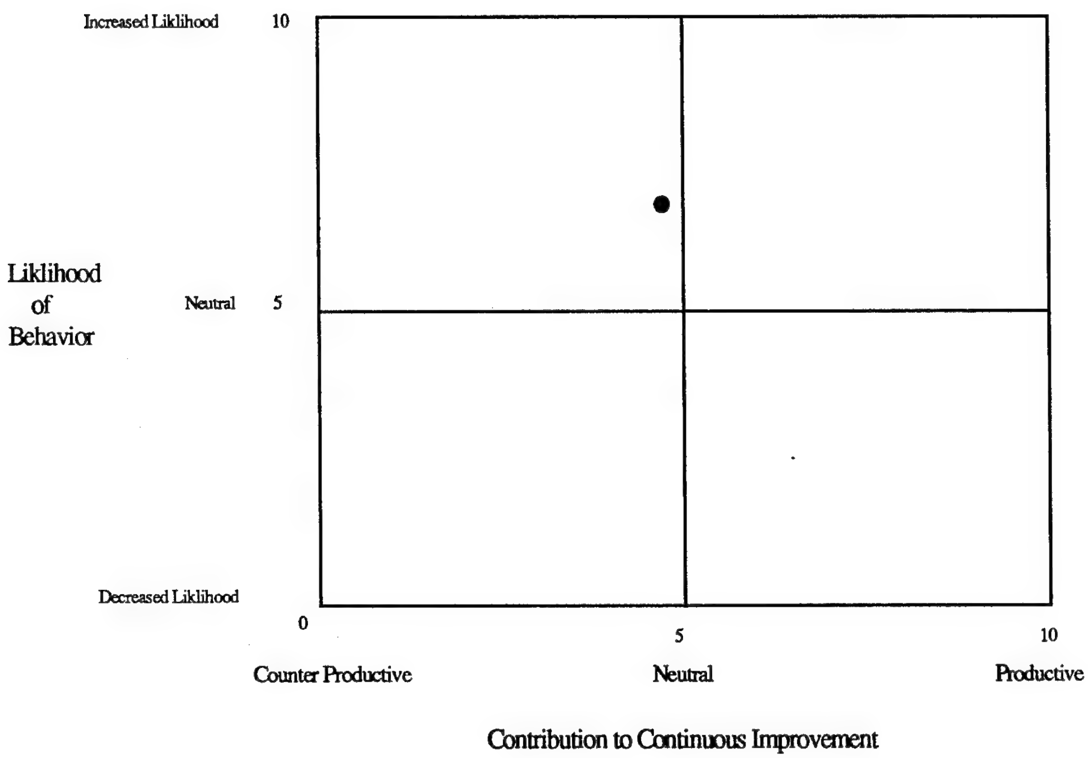

Figure A-42. Metric Three, Behavior Fifteen

Behavior 16. Overstating the budgeted cost.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
	Participant Ratings										
How well does the metric drive the behavior?	-	-	-	-	-	2	1	1	-	-	6.75
How well does the behavior contribute to CI?	-	2	1	1	-	-	-	-	-	-	2.75

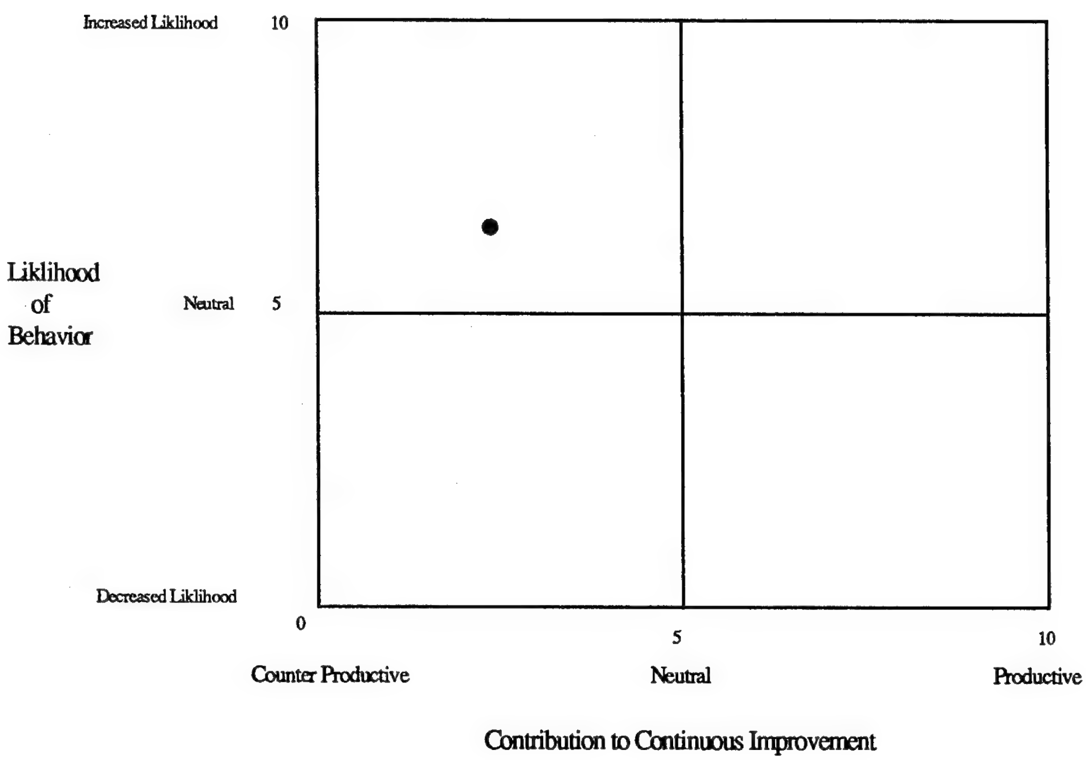

Figure A-43. Metric Three, Behavior Sixteen

Behavior 17. Causes cancellation of programs.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
	Participant Ratings										
How well does the metric drive the behavior?	-	-	-	-	1	1	1	-	1	-	6.75
How well does the behavior contribute to CI?	-	-	-	2	1	-	1	-	-	-	5.00

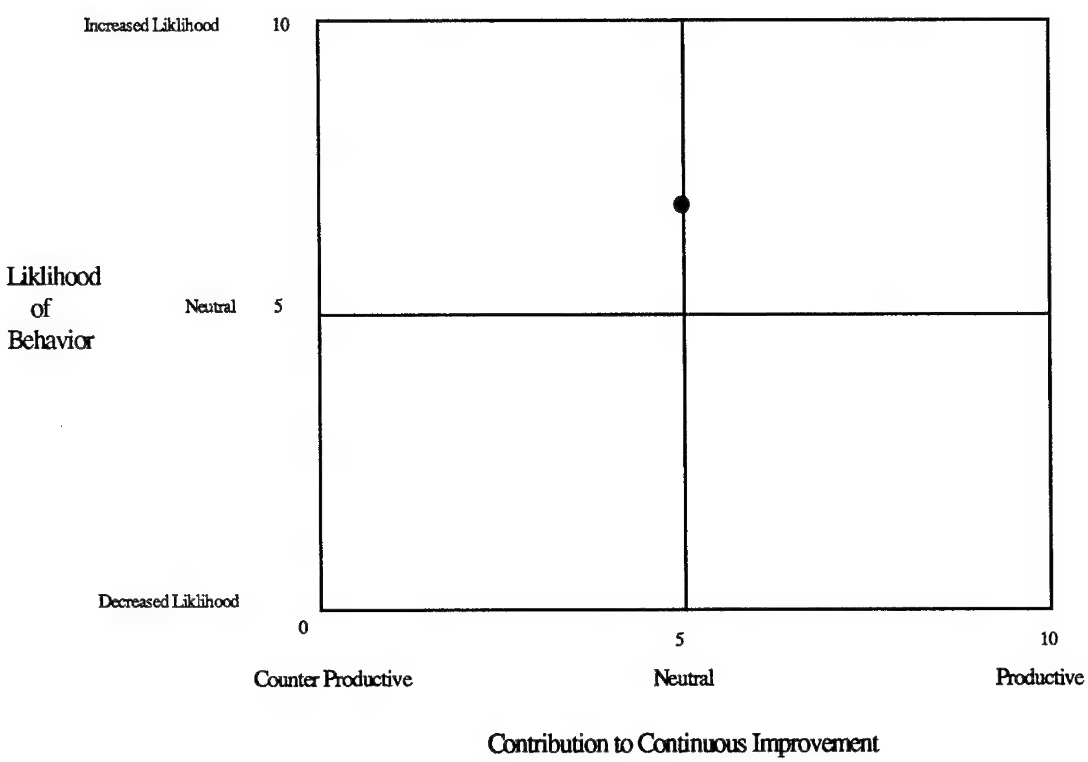

Figure A-44. Metric Three, Behavior Seventeen

Behavior 18. Helps generate new cost estimating techniques.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	2	2	-	-	-	6.50
How well does the behavior contribute to CI?	-	-	-	-	1	-	2	-	1	-	7.00

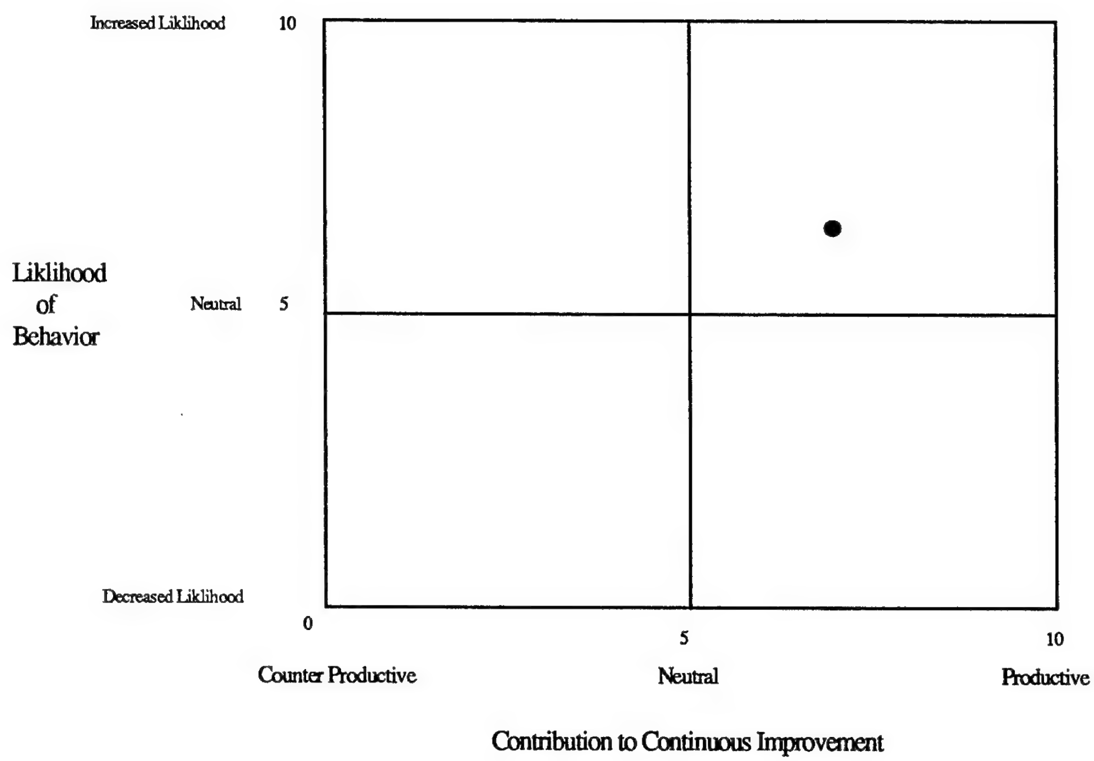

Figure A-45. Metric Three, Behavior Eighteen

Behavior 19. Causes the contractor to cut corners.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	1	1	1	-	-	6.50
How well does the behavior contribute to CI?	1	3	-	-	-	-	-	-	-	-	1.75

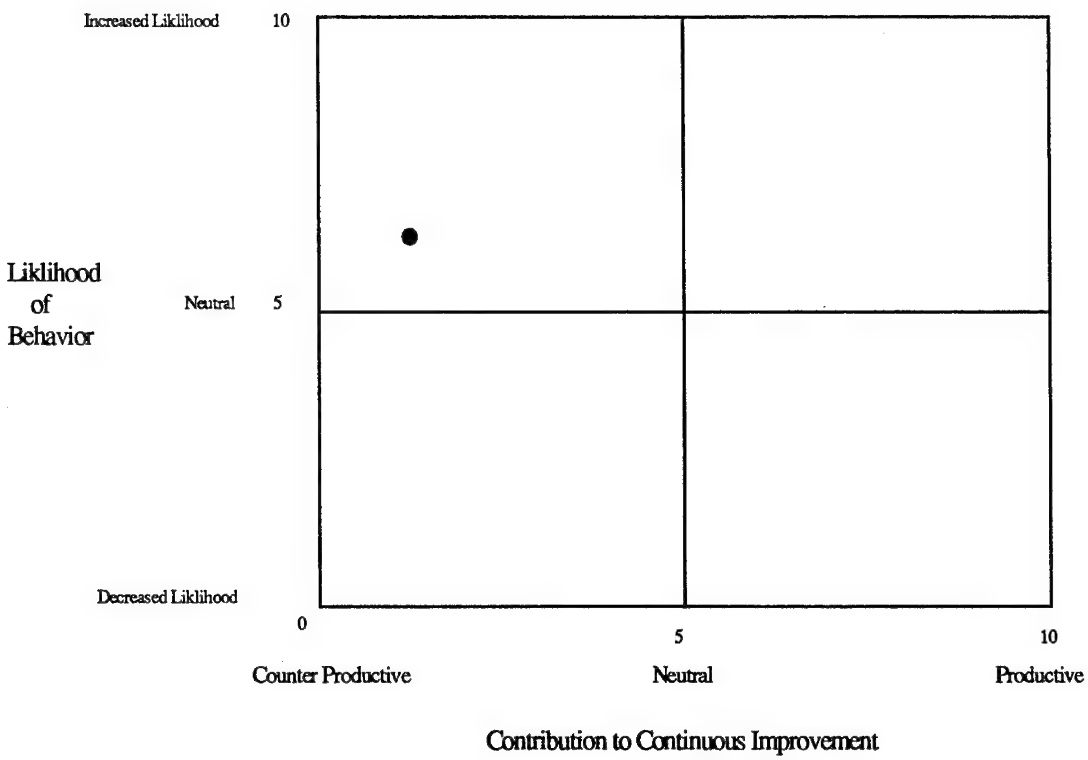

Figure A-46. Metric Three, Behavior Nineteen

Behavior 20. Report contractor performance in CPAS.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	3	-	-	-	-	1	6.25
How well does the behavior contribute to CI?	-	-	-	-	2	-	-	1	1	-	6.75

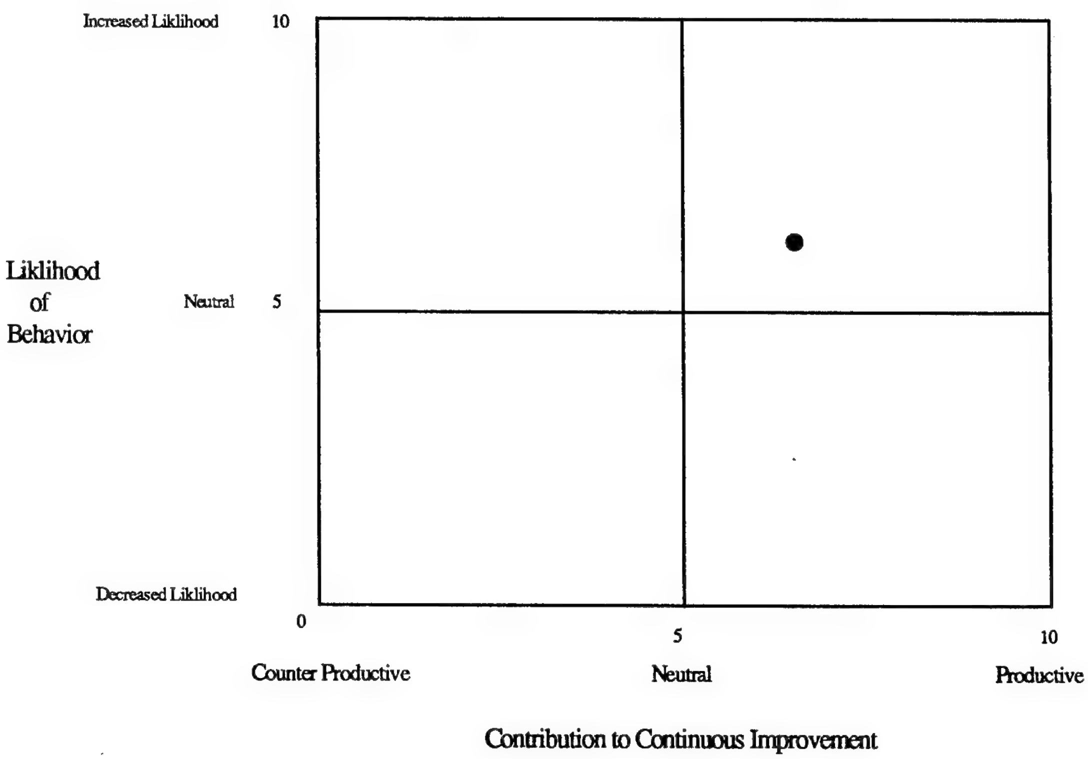

Figure A-47. Metric Three, Behavior Twenty

Behavior 21. Indicates the need for better cost estimating techniques.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	1	1	1	1	-	-	-	5.50
How well does the behavior contribute to CI?	-	-	-	-	-	1	1	1	1	-	7.50

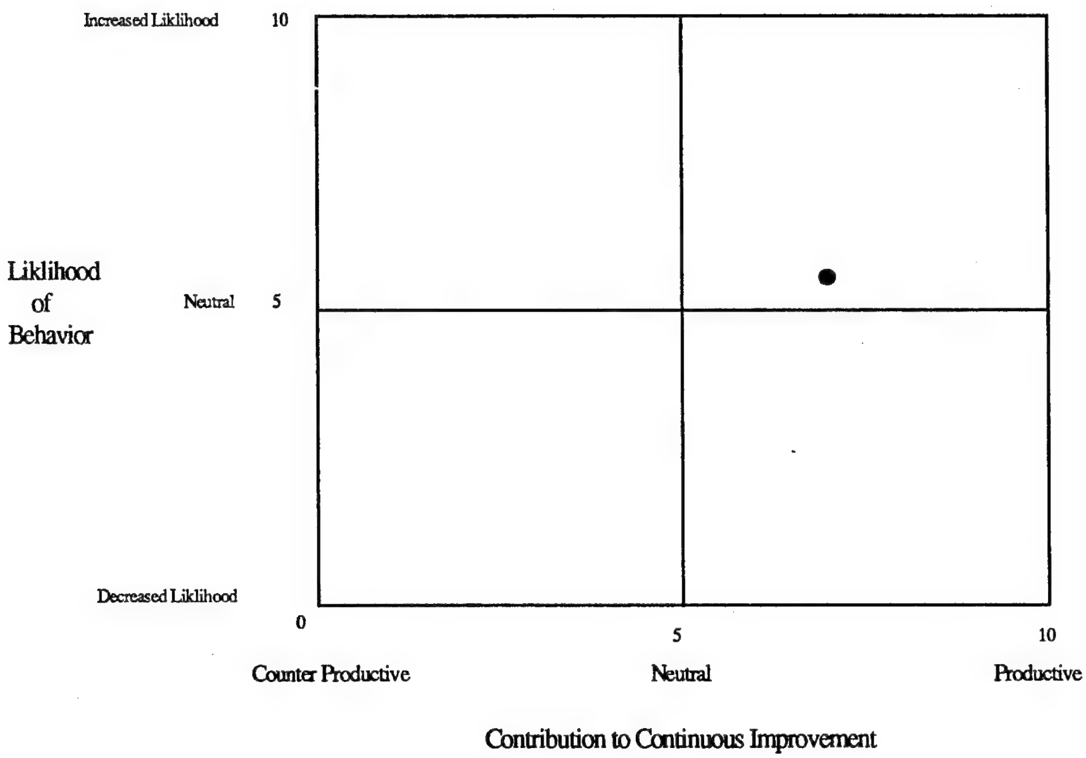

Figure A-48. Metric Three, Behavior Twenty-one

Behavior 22. Biased methods for contractor taking credit for work performed.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	1	-	1	1	1	-	-	-	5.25
How well does the behavior contribute to CI?	2	-	2	-	-	-	-	-	-	-	2.00

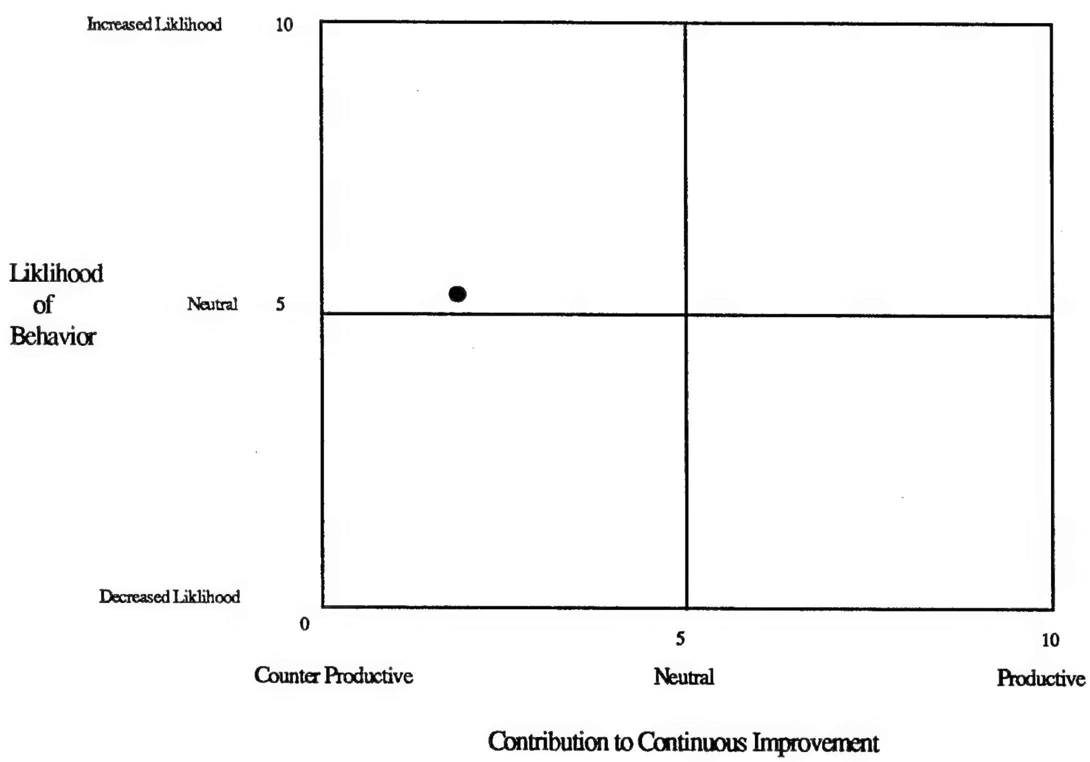

Figure A-49. Metric Three, Behavior Twenty-two

Metric 4: Program Cost Disconnect

Metric 4 Behavior 1. Add or delete content, speed or slow schedule, as required.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
	Participant Ratings										
How well does the metric drive the behavior?	-	-	-	-	-	-	-	2	-	2	9.00
How well does the behavior contribute to CI?	-	-	1	2	-	-	1	-	-	-	4.50

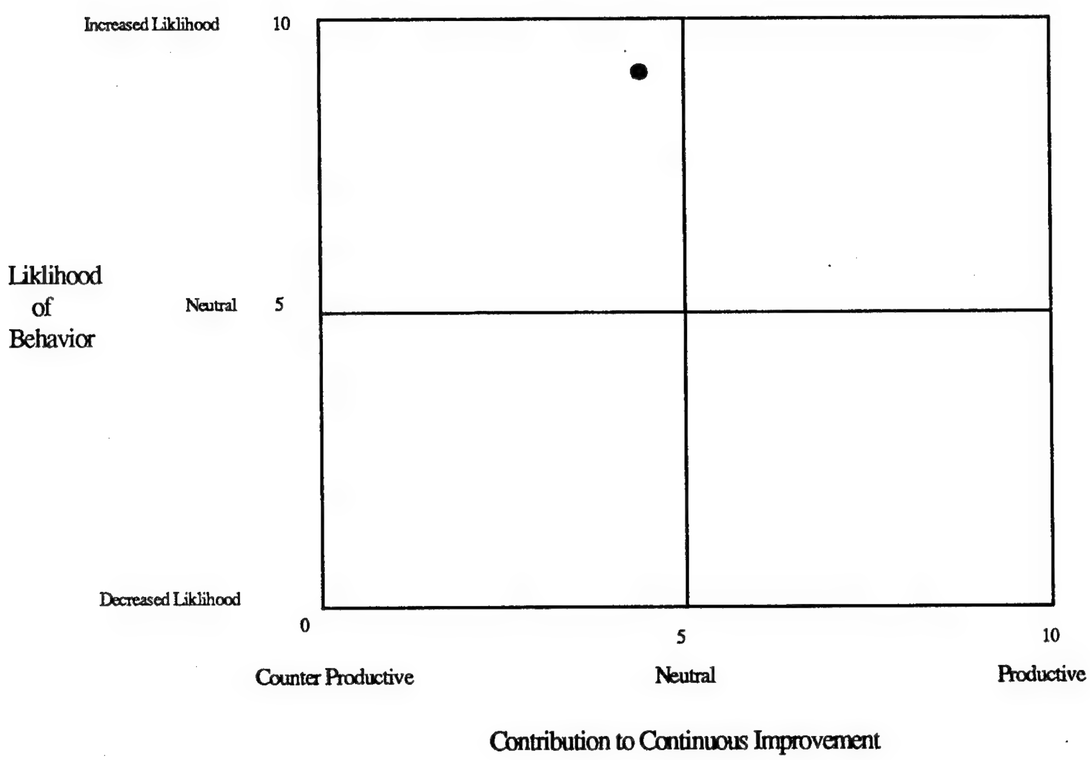

Figure A-50. Metric Four, Behavior One

Behavior 2. Restructure program plan to match budget.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	1	1	2	-	8.25
How well does the behavior contribute to CI?	-	-	1	2	-	-	1	-	-	-	4.50

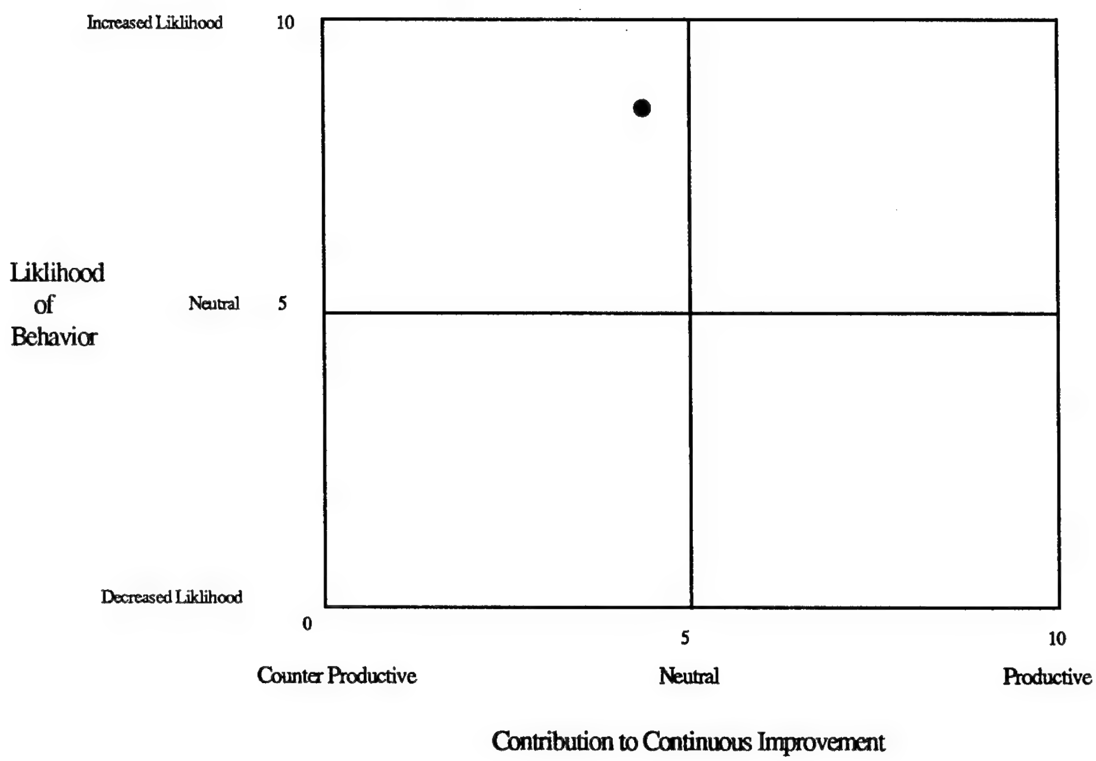

Figure A-51. Metric Four, Behavior Two

Behavior 3. Causes SPO to carefully plan dev., production, and O&M funding.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	-	2	1	1	-	7.75
How well does the behavior contribute to CI?	-	-	-	-	-	-	-	1	3	-	8.75

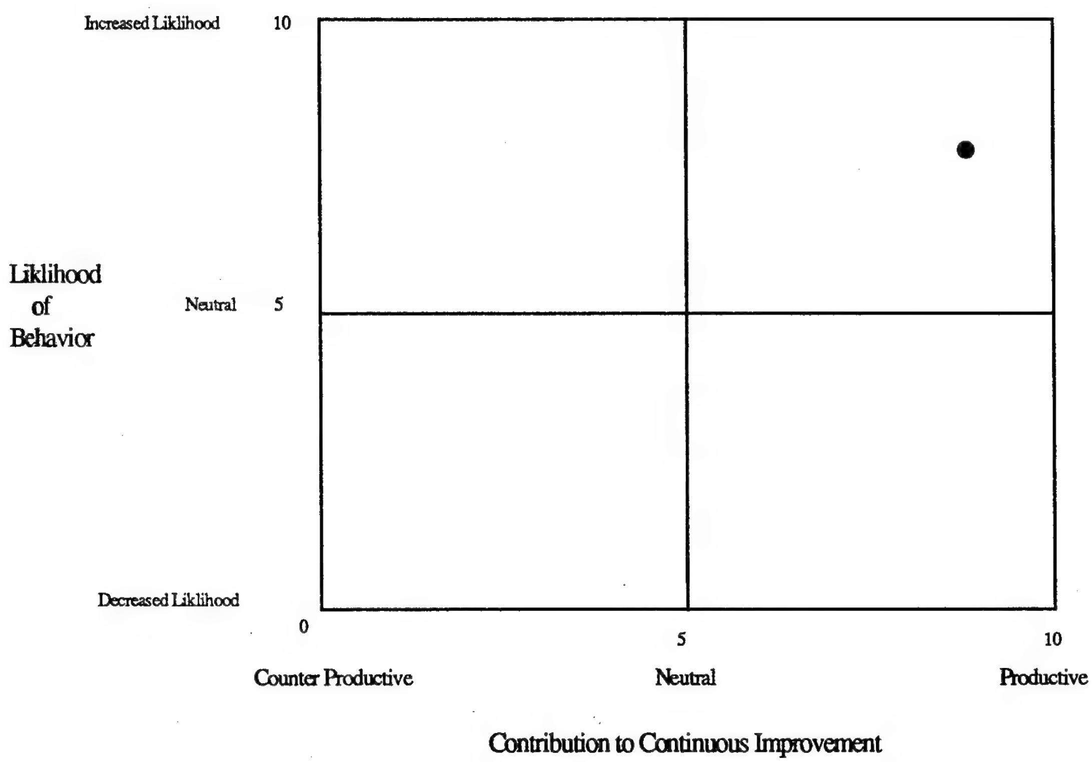

Figure A-52. Metric Four, Behavior Three

Behavior 4. Causes speculation on political environment--wasted effort.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	-	1	-	2	-	7.50
How well does the behavior contribute to CI?	-	2	1	1	-	-	-	-	-	-	2.75

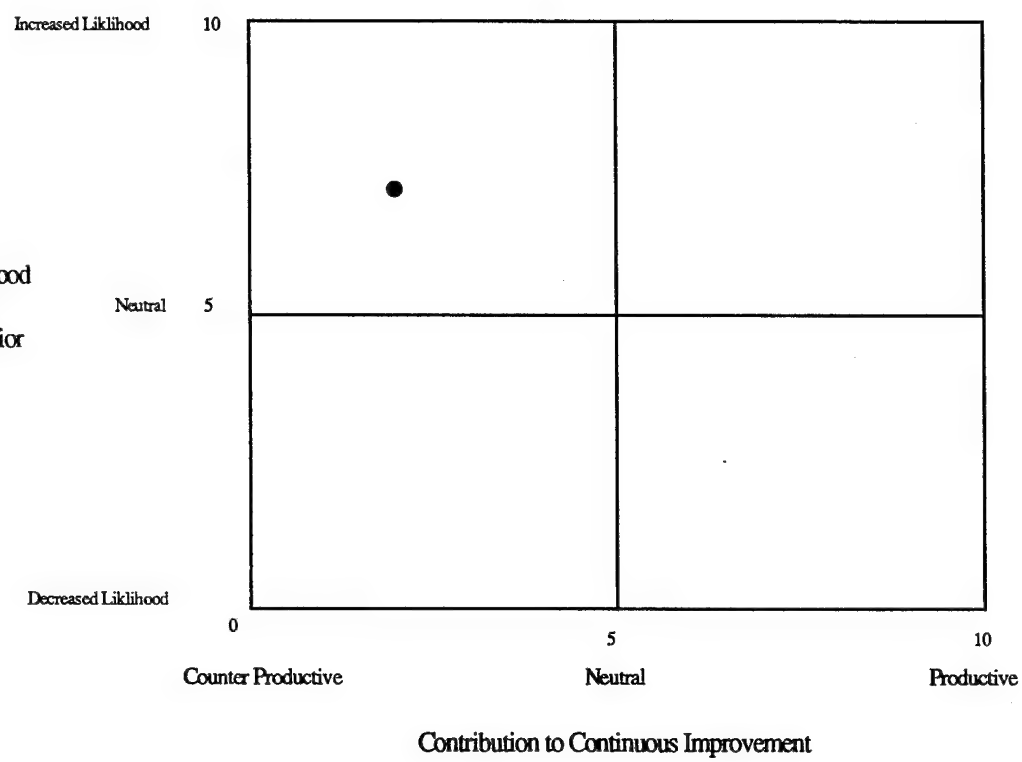

Figure A-53. Metric Four, Behavior Four

Behavior 5. Causes the SPO to request more funds.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
	Participant Ratings										
How well does the metric drive the behavior?	-	-	-	-	-	2	1	-	1	-	7.00
How well does the behavior contribute to CI?	-	1	1	-	1	-	1	-	-	-	4.25

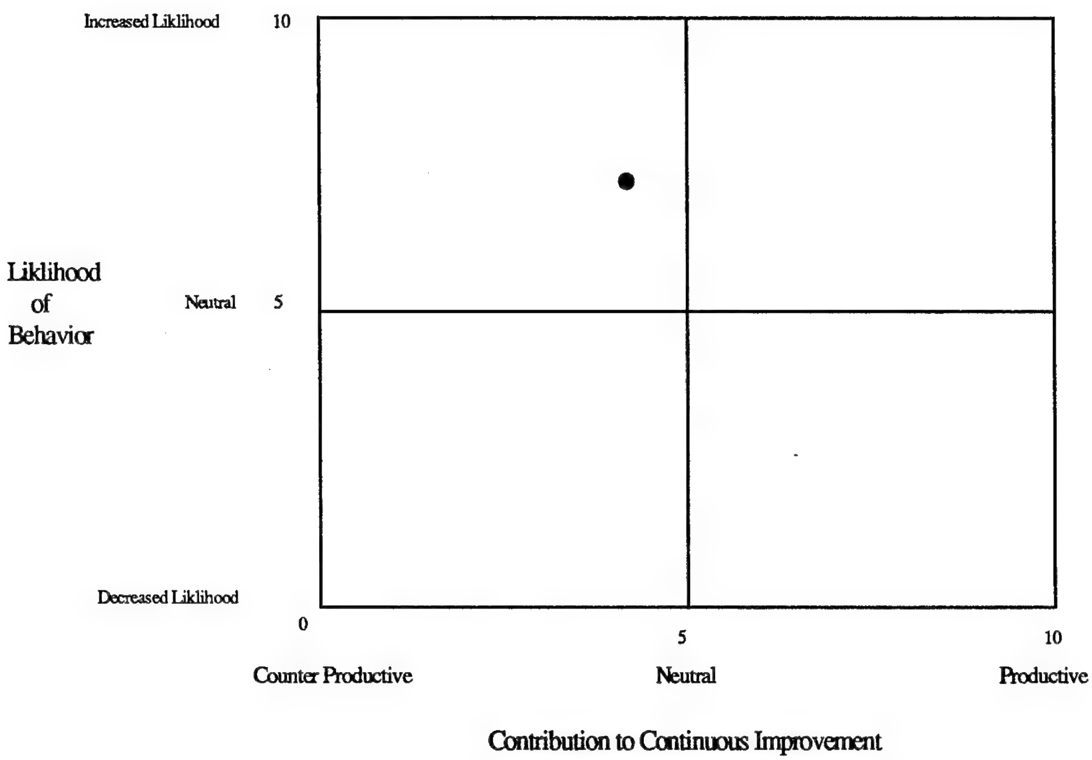

Figure A-54. Metric Four, Behavior Five

Behavior 6. Provides justification for increased funding.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	1	1	-	1	1	-	7.00
How well does the behavior contribute to CI?	-	-	-	-	1	1	1	1	-	-	6.50

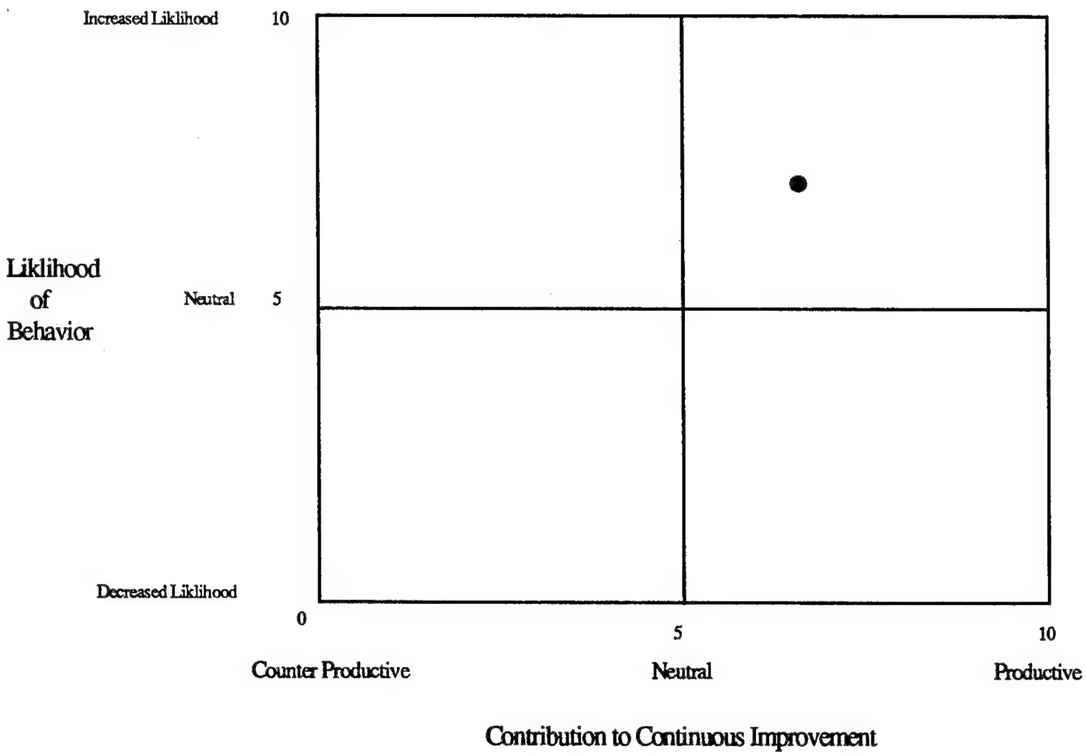

Figure A-55. Metric Four, Behavior Six

Behavior 7. Causes the SPO to request reprogramming of money.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	-	2	1	1	-	-	6.75
How well does the behavior contribute to CI?	-	-	-	1	1	1	1	-	-	-	5.50

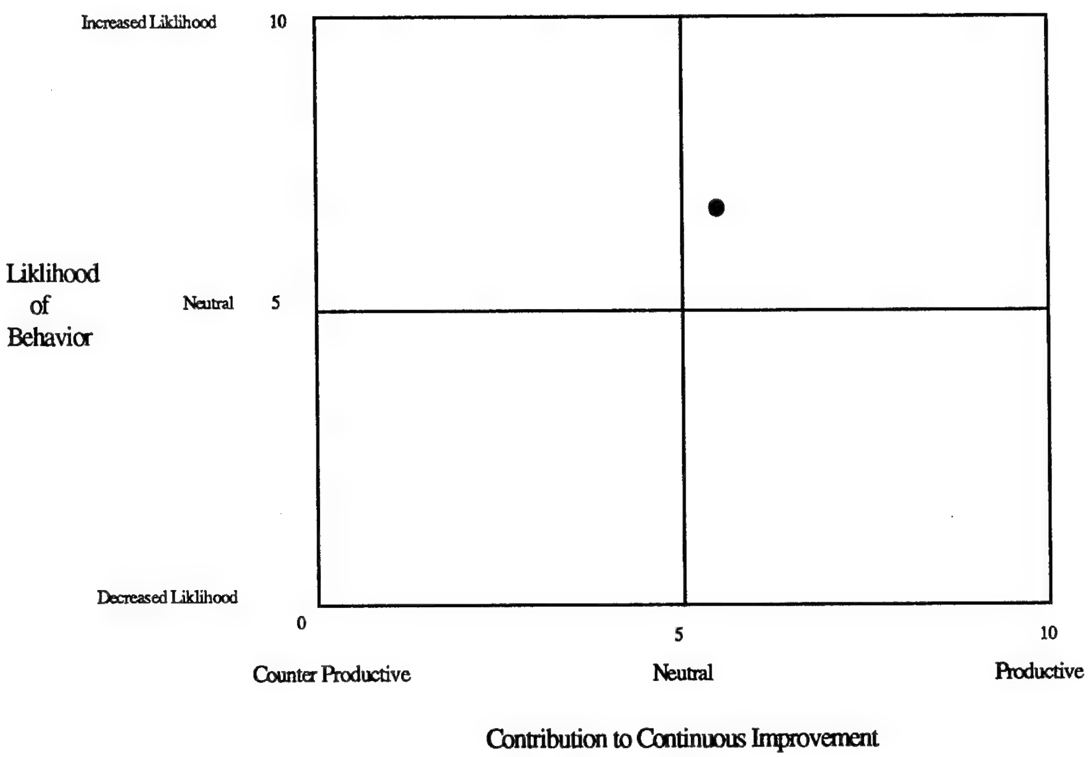

Figure A-56. Metric Four, Behavior Seven

Behavior 8. Cause revisions to requirements process.

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	-	2	1	-	1	-	-	6.00
How well does the behavior contribute to CI?	-	-	-	-	1	-	-	2	1	-	7.50

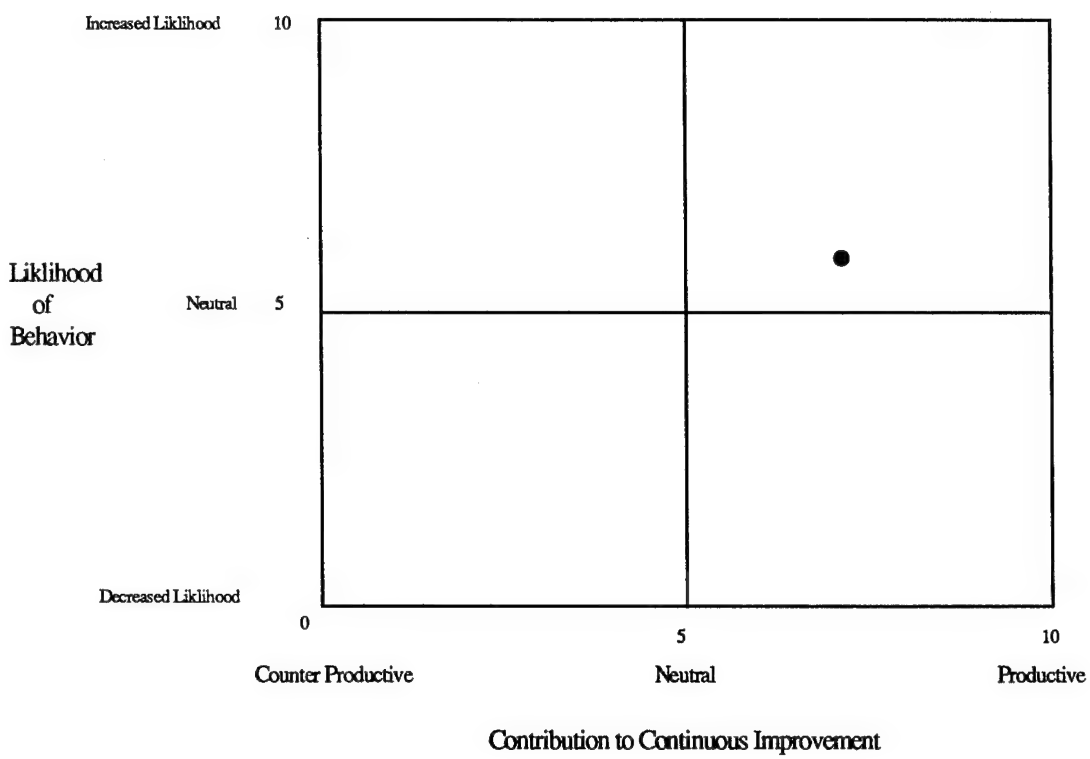

Figure A-57. Metric Four, Behavior Eight

Behavior 9. Causes reduction in number of programs (fund fewer better).

** Criteria**	Participant Ratings										
	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	1	-	1	2	-	-	-	-	5.00
How well does the behavior contribute to CI?	-	-	-	-	-	4	-	-	-	-	6.00

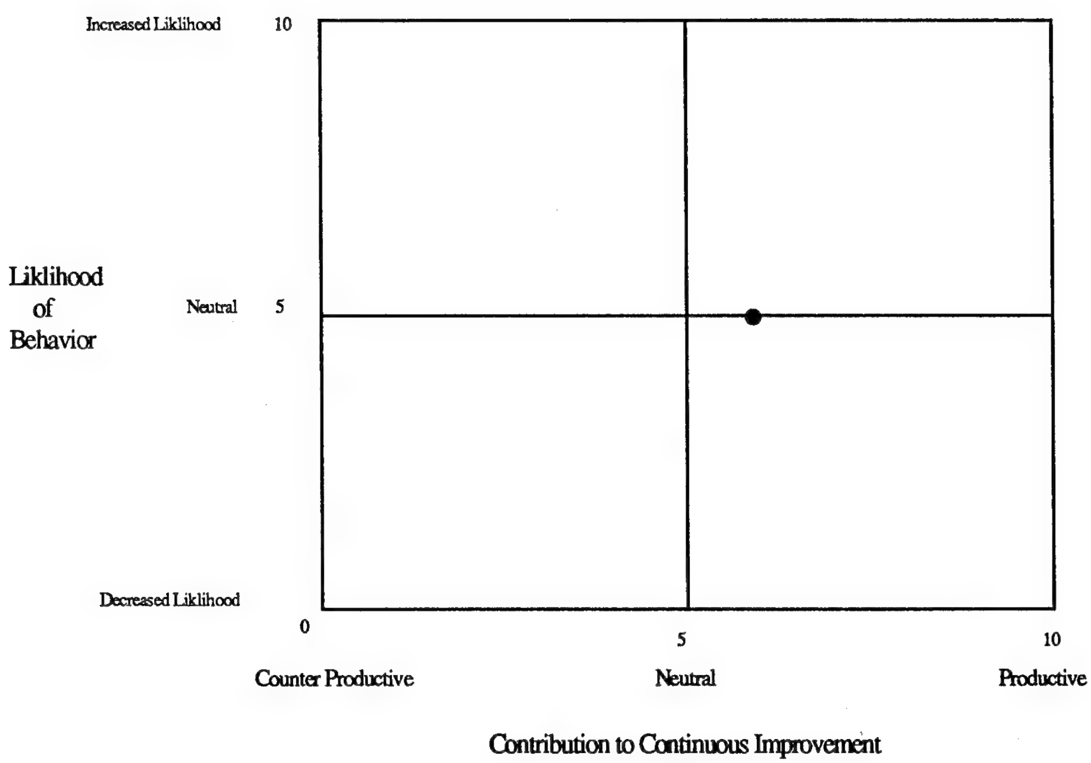

Figure A-58. Metric Four, Behavior Nine

Behavior 10. Request direction for use of excess funds.

** Criteria**	1	2	3	4	5	6	7	8	9	10	Mean
How well does the metric drive the behavior?	-	-	-	1	3	-	-	-	-	-	4.75
How well does the behavior contribute to CI?	-	-	-	-	1	1	-	2	-	-	6.75

Participant Ratings

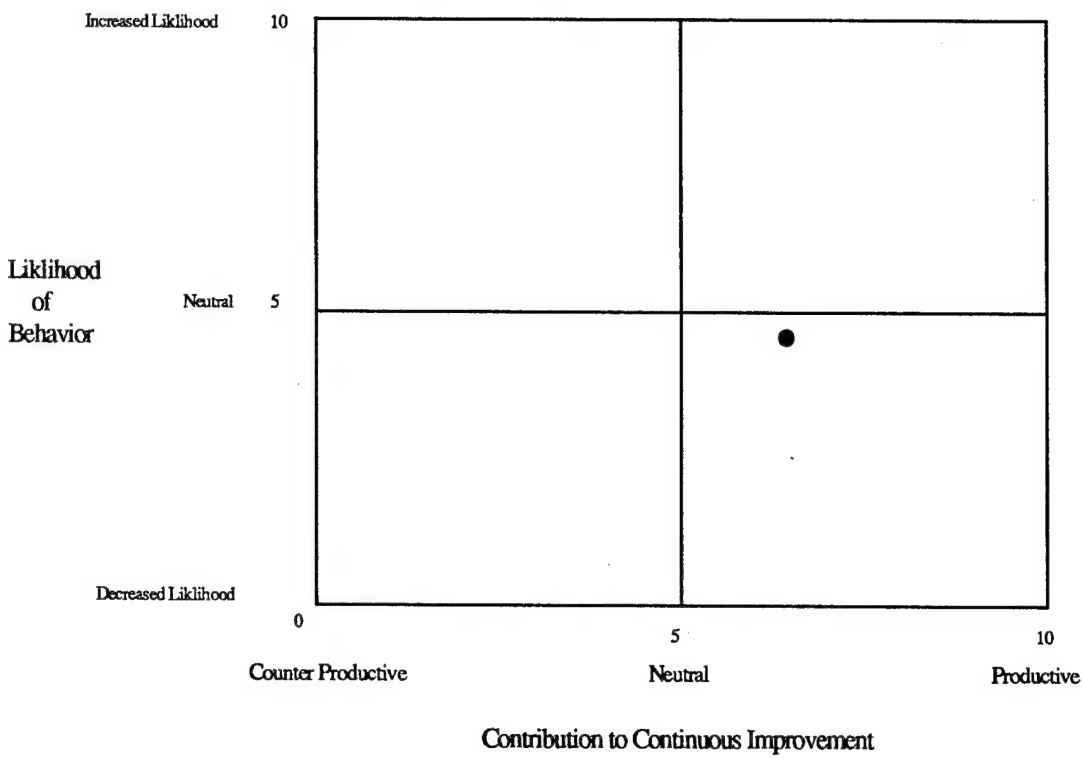

Figure A-59. Metric Four, Behavior Ten

Bibliography

Adams, Scott M., Joseph Sarkis, and Don Lies. "The Development of Strategic Performance Metrics," Engineering Management Journal 7: 24-31 (1 March 1995).

Air Force Materiel Command. The Metrics Handbook: Quality and Reliability Assurance. Wright-Patterson AFB OH: HQ AFMC, 31 August 1993.

Air Force Systems Command. The Metrics Handbook. Andrews AFB MD: HQ AFSC, August 1991.

Bowles, Jerry. "Is American Management Really Committed to Quality," Management Review: 42-46 (April 1992).

Crosby, Philip B. Quality is Free: The Art of Making Quality Certain. New York: McGraw-Hill Book Company, 1979.

Deming, W. Edwards. Out of the Crisis. Cambridge MA: Massachusetts Institute of Technology Center for Advanced Engineering Study, 1986

DoD Comptroller. "Key Criteria for Performance Measurement," DoD guideline. Washington DC. The Government Printing Office, 25 October 1992.

Dorrell, Msgt Bruce, 88th Communications Group. Personal interview. Wright-Patterson AFB OH, 04 April 1995.

Emmelhainz, L.W. "TQM Principles and Measures: Key to Successful Implementation," Air Force Journal of Logistics 15: 34-37 (Summer 1991).

Federal Quality Institute. "Keep Up With How Quality Management Ties in With Reinventing Government," Federal Quality News: Telling the Quality Story in Government 22: 4 (December 1993/January 1994).

Flynn, Barbara B. "Managing for Quality in the U.S. and in Japan," Interfaces, 22: 69-80 (September-October 1992).

Garvin, David A. Managing Quality: The Strategic and Competitive Edge. New York: The Free Press, 1988.

Grant, Kevin P. and Wendell P. Simpson. "The Metric Diagnostic Chart: A Metric for Metrics". Proceedings: PMI'93 Seminar/Symposium Proceedings. 1993.

Hare, Paul A. Creativity in Small Groups. Beverly Hills CA: Sage Publications, 1982.

Hayes, Robert J. and Lawrence M. Miller. An Evaluation of Schedule Metrics Used Within Aeronautical Systems Center. MS Thesis, AFIT/GSM/LSY/92S-12. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1992 (AD-A260113).

Juran, Joseph M. "You Say You Want a Revolution," Journal of Business Strategy, 14: 27 (September-October 1993).

-----. Juran on Leadership for Quality. New York: The Free Press, 1989.

Lyman, Shawn P. Application of the Deming Management Method to Implementing Total Quality in the DoD. MS Thesis, AFIT/GLM/LSQ/91S-43. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1991 (AD-A246740).

Moen, Captain Ken, Armstrong Laboratory. Personal interview. AL/HRGA, Wright-Patterson AFB OH, 24 April 1995.

Oakland, John S. Total Quality Management. Oxford: Heinemann Professional Publishing, 1989.

Tayntor, Christine B. "Partners in Excellence: Metrics and Productivity Programs," Information Systems Management: 81-83 (Winter 1994).

Virginia Productivity Center. Managing Quality and Productivity in Aerospace and Defense. Contract No. MDA903-85-C-0237 Managed by Defense Systems Management College, Fort Belvoir VA, November 1989.

Walton, Mary. The Deming Management Method. New York: Dodd, Mead, & Company, 1986.

Vita

Kevin Byrne was born in Seattle, Washington on 17 November 1970. After graduation from Kennewick High School in 1989, Kevin attended the United States Air Force Academy Preparatory School. One year later he was accepted into the United States Air Force Academy where he graduated with a Bachelors of Science in Management. Following graduation from the academy in June, 1994, Kevin attended the Air Force Institute of Technology at Wright-Patterson AFB, OH. He is currently enrolled in the Graduate Systems Management (GSM) program where he will receive a Masters of Science in Systems Management. Following graduation in September, 1995, Kevin will assume his duties at Phillips Laboratory in Albuquerque, New Mexico.

P.O. Box 7272
Kennewick, WA 99336

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	SEPTEMBER 1995	Master's Thesis

4. TITLE AND SUBTITLE
AN EVALUATION OF COST METRICS USED WITHIN AERONAUTICAL SYSTEMS CENTER

5. FUNDING NUMBERS

6. AUTHOR(S)
Kevin M. Byrne, 2Lt, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
WPAFB, OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GSM/LA/95S-01

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Aeronautical Systems Center (ASC) system program offices rely on cost metrics to acquire information on program effectiveness and organizational improvement. This study was an attempt to determine the effectiveness of ASC's cost metrics. The effectiveness of the metrics were determined by the desirability of the behaviors which are exhibited when individuals utilize the specific metrics. Using an expert panel to evaluate the metrics, data was collected using the Group Research Laboratory for Logistics located in Armstrong Laboratory at Wright Patterson AFB, OH. Metric diagnostic charts were utilized to arrange the data into a format which would provide effectiveness information. The results of the study indicated that management should be aware that an integrated approach when utilizing metrics may provide supplementary information about a particular situation. Additionally, management should aim at finding ways to reduce the counter productive behaviors for each metric. Currently, existing metrics are excellent in increasing the likelihood of the behaviors. However, this research shows that a large number of counter productive behaviors are being driven. Finally, management needs to consider how they are communicating the definitions of the "common" cost metrics.

14. SUBJECT TERMS
Acquisition, Procurement, Measurement, Systems Management

15. NUMBER OF PAGES
111

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL